

SEARCH REQUEST FORM

50080

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: CS/Team. Tully
 Searcher Phone #: 308-45010
 Searcher Location: Biotec Lib
 Date Searcher Picked Up: 8/31/01
 Date Completed: 9/4/01
 Searcher Prep & Review Time: _____
 Clerical Prep Time: 4 min
 Online Time: 3 min

Type of Search

NA Sequence (#) _____
 AA Sequence (#) 1
 Structure (#) _____
 Bibliographic _____
 Litigation _____
 Fulltext _____
 Patent Family _____
 Other _____

Vendors and cost where applicable

STN _____
 Dialog _____
 Questel/Orbi: _____
 Dr.Link _____
 Lexis/Nexis _____
 Sequence Systems ABSSO1
 WWW/Internet _____
 Other (specify) _____

STIC-Biotech/ChemLib

50080

From: Jiang, Dong
Sent: Thursday, August 30, 2001 6:15 PM
To: STIC-Biotech/ChemLib
Subject: SN09/333,159

6/14/99
McCarthy

Please search SEQ ID NO:47 (back translation only)

-issued
-commercial

Please send results on paper to Dong Jiang in 10B-01 (mail stop CM1-10C01). Thank you very much for your help.

Dong Jiang (78243)
703-305-1345
U.S. Patent and Trademark Office
Art Unit 1646
dong.jiang@uspto.gov
CM1-10B01
Mail stop: CM1-10C01

RECEIVED
AUG 31 2001
STIC

OM of: US-09-333-159-47 to: GenEmbl.* out_format : pfs
 Date: Sep 1, 2001 3:21 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
 -MODEL-frame2_p2n.model -DEV-rlp
 -Q/cgnt2/USPTO_spool/US09333159/runat_31082001_122330_1384/app_query.fasta_1.488
 -DB-GenEmbl -OFFMT-fastap -SUFFIX-p2n.rge -GAPOP-12.000
 -GAPEXT-4.000 -MINMATCH-0.100 -LOOPEXT-0.000 -LOOPEXT-0.000
 -GAPOP-4.000 -GAPEXT-0.050 -XGAPOP-10.000 -XGAPEXT-0.500
 -FCAPOP-6.000 -FCAPEXT-7.000 -XGAPOP-10.000 -XGAPEXT-0.500
 -DELOP-6.000 -DELEXT-7.000 -START-1 -MATRIX-blosum62
 -TRANS-human40.cdi -LIST-45 -DOCALLIGN-200 -THR_SCORE-pct
 -THR_MAX-100 -THR_MIN-0 -ALIGN-15 -MODE-LOCAL -OUTFMT-pfs
 -NORM-ext -MINLEN-0 -MAXLEN-200000000
 -USER-US09333159_cGNTL_1_5294 -NCPU-6 -ICPU-3 -LONGLOG -NO_XLPHY
 -WAIT -THREADS-1

Search information block:

Query: US-09-333-159-47
 Query length: 423
 Database: GenEmbl.*
 Database sequences: 1344157
 Database length: -856060004
 Search time (sec): 1547.640000

score_list:	Strd Orig	ZScore	Escore	Len	Documentation
9b_pr9:HSAL	+ 1308.00	2043.50	1.4e-105	2626	X76488 H.sapiens mRNA for lys
9b_pr10:HSAL	+ 1306.00	2040.82	1.9e-105	2493	M74775 Human lysosomal acid li
9b_pr9:HSAL	+ 1298.00	2028.24	9.8e-105	2481	Z31690 H.sapiens (HepG2) LAL m
9b_pr10:HS008464	+ 1298.00	2028.24	9.8e-105	2481	U08464 Human lysosomal acid li
9b_pr10:HS1497	+ 1220.50	1903.81	8.3e-98	3144	S81497 lysosomal acid lipase-1
9b_pr1:MMLVACLY	+ 1206.00	1893.53	1.1e-96	2358	Z31689 M.musculus (C57 Black/6
9b_pr1:MMLVACLY	+ 1167.00	1826.83	1.6e-93	1378	A26689 Precursor of rabbit gast
9b_pr1:MMLVACLY	+ 1167.00	1826.83	1.6e-93	1378	A26690 Precursor of rabbit gast
9b_pr1:MMLVACLY	+ 1166.50	1826.21	1.7e-93	1336	A01157 R.norvegicus mRNA for p
9b_pr1:MMLVACLY	+ 1154.50	1804.70	2.8e-92	1651	Y13899 Canis familiaris mRNA f
9b_pr1:MMLVACLY	+ 1141.50	1788.35	2.2e-91	1137	AR039022 Sequence 2 from paten
9b_pr1:MMLVACLY	+ 1141.50	1788.33	2.2e-91	1137	AR039023 Sequence 2 from paten
9b_pr1:MMLVACLY	+ 1141.50	1788.33	2.2e-91	1140	AR039023 Sequence 4 from paten
9b_pr1:MMLVACLY	+ 1141.50	1788.33	2.2e-91	1140	AR039023 Sequence 4 from paten
9b_pr1:MMLVACLY	+ 1141.50	1788.28	2.3e-91	1146	AR039024 Sequence 6 from paten
9b_pr1:MMLVACLY	+ 1141.50	1788.28	2.3e-91	1146	AR039024 Sequence 6 from paten
9b_pr1:MMLVACLY	+ 1141.50	1787.88	2.4e-91	1198	A57760 Sequence 5 from paten
9b_pr1:MMLVACLY	+ 1141.50	1785.68	3.2e-91	1528	A57756 Sequence 1 from Paten
9b_pr1:MMLVACLY	+ 1141.50	1785.68	3.2e-91	1528	AR039021 Sequence 1 from paten
9b_pr1:MMLVACLY	+ 1141.50	1785.68	3.2e-91	1528	AR039021 Sequence 1 from paten
9b_pr1:MMLVACLY	+ 1139.00	1782.41	4.8e-91	1417	L26319 Bos taurus pregastric e
9b_pr1:MMLVACLY	+ 1132.50	1774.15	1.4e-90	1137	A39301 Sequence 2 from Paten
9b_pr1:MMLVACLY	+ 1132.50	1774.12	1.4e-90	1140	A39303 Sequence 4 from Paten
9b_pr1:MMLVACLY	+ 1132.50	1774.08	1.4e-90	1146	A39305 Sequence 1 from Paten
9b_pr1:MMLVACLY	+ 1132.50	1771.48	2.0e-90	1528	A39300 Sequence 1 from Paten
9b_pr1:MMLVACLY	+ 1124.00	1759.07	9.6e-90	1365	A50997 Human mRNA for gastric
9b_pr1:MMLVACLY	+ 1124.00	1759.07	9.6e-90	1367	A01046 H.sapiens mRNA for huma
9b_pr1:MMLVACLY	+ 1124.00	1759.07	9.6e-90	1367	A12714 Pregastric lipase. 8/19
9b_pr1:MMLVACLY	+ 1003.00	1570.55	3.0e-79	1048	A57758 Sequence 3 from Paten
9b_pr1:MMLVACLY	+ 695.00	1082.55	4.6e-52	1308	AX068259 Sequence 21 from Paten
9b_pr1:MMLVACLY	+ 691.00	1081.90	5.0e-52	699	AX068249 Sequence 11 from Paten
9b_pr1:MMLVACLY	+ 675.00	1008.19	6.4e-48	148720	AC019353 Homo sapiens clone
9b_pr1:MMLVACLY	+ 675.00	1007.63	6.9e-48	158195	AC015461 Homo sapiens clone
9b_pr1:MMLVACLY	+ 675.00	1006.83	7.6e-48	172834	AC011277 Homo sapiens clone
9b_pr1:MMLVACLY	+ 675.00	1006.41	8.1e-48	181170	AL353113 Homo sapiens chrom
9b_pr1:MMLVACLY	+ 665.00	1035.91	1.8e-49	1211	Y14367 Drosophila melanogaster
9b_pr1:MMLVACLY	+ 658.00	994.04	3.9e-47	36616	U70780 Caenorhabditis elegans
9b_pr1:MMLVACLY	+ 638.50	962.74	2.2e-45	38822	AC084473 Caenorhabditis brig
9b_pr1:MMLVACLY	+ 637.00	960.82	2.8e-45	36982	U64849 Caenorhabditis elegans

9b_in1:AC004766 + 636.00 950.81 1.0e-44 93932 AC004766 Drosophila mela
 9b_in1:AC008342 - 636.00 945.18 2.1e-44 175007 AC008342 Drosophila mela
 9b_in1:AC017903 - 636.00 943.46 2.6e-44 211784 AC017903 Drosophila mela
 9b_in2:AE003811 - 636.00 941.62 3.3e-44 259445 AE003811 Drosophila mela
 9b_in2:AC013103 - 635.00 961.05 2.7e-45 25416 AC013100 Drosophila mela
 seq_name: gb_pr9:HSAL
 seq_documentation_block:
 LOCUS HSLAL 2626 bp mRNA PRI 25-FEB-1994
 DEFINITION H.sapiens mRNA for lysosomal acid lipase.
 ACCESSION X76488
 VERSION X76488.1 GI:434305
 KEYWORDS lysosomal acid lipase.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2626)
 AUTHORS Ameis,D., Merkel,M., Eckerskorn,C. and Greten,H.
 TITLE Purification, characterization and molecular cloning of human
 hepatic lysosomal acid lipase
 JOURNAL Eur. J. Biochem. 219 (3), 905-914 (1994)
 MEDLINE 94155897
 REFERENCE 2 (bases 1 to 2626)
 AUTHORS Ameis,D.
 TITLE Direct Submission
 JOURNAL Submitted (29-NOV-1993) D. Ameis, Medical Department, University
 Hospital Eppendorf, Martinistrasse 52, 20246 Hamburg, FRG
 FEATURES
 Location/Qualifiers
 source
 1..2626
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="10"
 /tissue_type="liver"
 146..1345
 /gene="LAL/LIPA"
 146..1345
 /gene="LAL/LIPA"
 /EC_number="3.1.1.13"
 /note="sterol esterase"
 /codon_start=1
 /evidence=experimental
 /product="lysosomal acid lipase"
 /protein_id="CAAS4026.1"
 /db_xref="GI:434306"
 /translation="MKRFLGLVCLVLTLSHSGSGKLTAVDPETNMNYSSEIYSW
 GF7SEEVLTEDGYILCLNRIHPGRKNHSDKPKPVVFLQHLGLADSSNNVTNLANS
 LPHLRLKDFGDEFLPQSAFLKWLGVCTHVLKELCGNLCLLCGFGFNRNLNMSR
 SLGFLADAGFDVMSNGRNTWSRKHTLSVSDSEFAFSYDEMAKYDLPLASINFL
 NKTGQSOVYTVGHSQGTIGFIAPSQIPELAKRMFKFALGVASVAFCTSPMAKLR
 VDYTHSPAGTSVQNNLHMSQAVKQFQAFDMGSSAKNYFHYNQSPYPTYNVMDL
 VPTAVNSGGHDLADYVDVNVILLTQITNLVPHESIPWEHLDFIWLGLDAPWRLYNKII
 NLRKQY"
 146..226
 /gene="LAL/LIPA"
 /evidence=experimental
 227..373
 /gene="LAL/LIPA"
 /note="propeptide"
 /evidence=experimental
 374..1342
 /gene="LAL/LIPA"
 /evidence=experimental
 /product="lysosomal acid lipase"
 741 a 527 c 547 g 811 t
 BASE COUNT
 ORIGIN
 alignment_scores:
 Quality: 1308.00 Length: 410
 Ratio: 3.836 Gaps: 2

[illegible]

```

seq_documentation_block:
LOCUS      HSLYACLY      2481 bp      mRNA      PRI      17-FEB-1997
DEFINITION H.sapiens (HepG2) LAL mRNA for lysosomal acid lipase.
ACCESSION  Z31650
VERSION    Z31690.1 GI:506430

```



```

925 CAGCGCTGTTAAATCCAAAAGTTTCAAGCCTTTGACTGGGGAAGCAGTGC 974
326 rLysAsnLeuGluLysCysAsnGlnProThrProValaGtTyrArgVala 343
:||||| :||||| :||| :||| :||| :||| :||| :||| :||| :|||
975 CAGAATTATTTTCATTACACAGAGTATCTCCACATACAAATGTGA 1024
343 rGAspMetThrValProThrAlaMetTyrPheGlyGlnAspTyrLeu 359
:||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1025 AGGACATGCTGTGCCAGCTGCAGTCTGGAGCGGGGTCAGACTGGCTT 1074
360 SerAsnProGluAspValLysMetLeuLeuSerGluValThrAsnLeu 376
:||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1075 GCAGATGCTCTACGACGCTCAATATCTTACTGACTCAGATCACCACCT 1124
376 eTyrHisLysAsnIleProGluTyrPheAlaHisValAspPheLeu 393
:||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1125 GTTCCATGAGAGCATTCGGAATGGGAGCATCTTGACTTCATTTGGGGCC 1174
393 euAspAlaProHisArgMetTyrAsnGluIleLeuHisLeuMetGlnGln 409
1175 TGGATGCCCTTGGAGGCTTATATAAATTAATTAATCTAATGAGGAAA 1224

seq_name: gb_pr10:HSU08464

seq_documentation_block:
LOCUS HSU08464 2481 bp mRNA PRI 23-JUN-1994
DEFINITION Human lysosomal acid lipase mRNA, complete cds.
ACCESSION U08464
VERSION U08464.1 GI:505052
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2481)
AUTHORS Du.H. and Gregory,G.A.
TITLE Structural Conservation of Putative Functional Motifs between Mouse
and Human Lysosomal Acid Lipase
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 2481)
AUTHORS Du.H.
TITLE Direct Submission
JOURNAL Submitted (08-Apr-1994) Hong Du, Division of Human Genetics,
Children's Hospital Medical Center, 3333 Bernet Street, Cincinnati,
OH 45229-3039, USA

FEATURES
source
1..2481
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="clone hLAL"
/clone_lib="HepG2 cDNA library/Lambda Zap"
/cell_line="HepG2"
/tissue_type="hepatoma"
1..33
/location="Qualifiers"
5'UTR 34..96
sig_peptide 34..1233
CDS
/product="lysosomal acid lipase"
/protein_id="AAB60328.1"
/db_xref="GI:505053"
/translation="MKRFLGLVCLVLTWLTLSBGRKLTAVDPETNMNVSEIISYW
GFPEEYLVEREDGYILCNRPKRHSDKPKPVFLQHGLLADSSNWTNLANS
SLGTLADAGFDVWNGSRNTRSKRHTLSVQDFWAFSDYEMAKYDLPASINFIL
NKTQEQVYVGHSGQTTIGFIAPSQIPELAKRMFEALGFVASVAFCTSPMAKLR
LPDLHLKIDFGDKFELQSAFLKWLHVCTHVKILKELCNLFLLCGFNERNLNMSR
VDVYTHSPACTSVQNMHLHWSQAVKFOAFDNGSSAKNFHYNQSPPTVYKDL
VPTAVWSSGHDLADVDVYNILLTQITNLNVHESIPWEHLDFINGLDAPRLINKII
NLMKRYQ"
79
variation
97..1230
mat_peptide
/product="lysosomal acid lipase"

```

```

variation
100
/replace="g"
1234..2480
polyA_signal 2459..2464
polyA_site 2480
BASE COUNT 692 a 477 c 513 g 799 t
ORIGIN
alignment_scores:
Quality: 1298.00 Length: 400
Ratio: 3.875 Gaps: 1
Percent Similarity: 83.750 Percent Identity: 61.000
alignment_block:
US-09-333-159-47 x HSU08464
Align seg 1/1 to: HSU08464 from: 1 to: 2481
12 SerHisArgMetGluMetTyrPheLeuIleLeuValaTyrMetPheG 28
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 AGTCCAGAAATGAAATGCCGTTCTTGGGGTGGTCTCTTTGGTTCT 74
28 nArgAsnValAsnSerValHisMetProThrLys.....AlaValAsp 43
:||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
75 CTGGACCTCTCATCTGAGGGTCTAGAGGAACTACAGCTGTGGATC 124
43 roGluAlaPheMetAsnIleSerGluIleLeuGlnHisGlnGlyTyrPro 59
:||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
125 CTGAACAACATGAATGTGAGTGAATTTATCTTTACTGGGATTCCT 174
60 CysGluGluTyrGluValAlaThrGluAspGlyTyrIleLeuSerValas 76
||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
175 AGTGAAGAAATACCTAGTTGACAGAGAGATGGATATATCTTGCCCTAA 224
76 nArgIleProArgGlyLeuValGlnProLysLysThrGlySerArgPro 93
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
225 CCGAATTCCTCATGGGAGGAGAACCATTCGCACAAAGGTCCTCAACAG 274
93 alValLeuLeuGlnHisGlyLeuValGlyGlyAlaSerAsnTrpIleSer 109
:||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
275 TTGCTCTTCCTGCAACATGGCTTCTGGCAGATTCTAGTAACTGGTCACA 324
110 AsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAspAlaGlyPheAs 126
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
325 AACCTTCCCAACAGACGCTGGGCTTCATCTTCTGCTGATGCTGTTTGA 374
126 pValTyrMetGlyAsnSerArgGlyAsnAlaTyrSerArgLysHisLys 143
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
375 CGTGTGGTGGGCAACAGACAGAGAAATACCTGCTCGGAACATAAGA 424
143 hrLeuSerIleAspGlnAspGluPheTrpAlaPheSerTyrAspGluMet 159
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
425 CACTCTCAGTTTCTCAGGATGAATCTGGGCTTTCAGTTATGATGATG 474
160 AlaArgPheAspLeuProAlaValIleAsnPheIleLeuGlnLysThrG 176
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
475 GCAAAATATGACCTACGAGCTTCATTAACCTTCACTCTGAATAAACTGG 524
176 yGlnGluLysIleTyrValGlyTyrSerGlnGlyThrThrMetGlyP 193
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
525 CCAAGAACAGTGTATTATGTGGTCTATTCTCAAGGCACCATATAGTT 574
193 heIleAlaPheSerThrMetProGluLeuAlaGlnLysIleLysMetTyr 209
:||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
575 TTATASCAITTTTCACAGATCCCTGAGCTGGCTAAAGATTAATATGTTT 624
210 PheAlaLeuAlaProIleAlaThrValLysHisAlaLysSerProGly 226
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
625 TTTGCTCTGGGCTCTGCTGGCTTCGCTGCTCTGTACTAGCCCTATGGC 674
226 rLysPheLeuLeuProAspMetMetIleLysGlyLeuPheGlyLysL 243
:||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

```



```

1201 AATACCACTGTAATGCTGGACCTTGAATACTTCCCTTCAGAGGGG 1245
seq_name: gb_pat1.A26689
seq_documentation_block: 1378 bp DNA PAT 05-APR-1995
LOCUS A26689 Precursor of rabbit gastric lipase coding sequence.
DEFINITION A26689
ACCESSION A26689
VERSION A26689.1 GI:905029
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
synthetic construct
artificial sequence.
REFERENCE
1 (bases 1 to 1378)
AUTHORS
Benicourt,C., Blanchard,C. and Junien,J.L.
TITLE
Recombinant gastric lipase from rabbit and pharmaceutical
compositions
JOURNAL
Patent: EP 0542629-A 9 19-MAY-1993;
INSTITUT DE RECHERCHE JOUVEINAL
FEATURES
Location/Qualifiers
source
1..1378
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 397 a 287 c 276 g 418 t
ORIGIN
alignment_scores:
Quality: 1167.00 Length: 402
Ratio: 3.505 Gaps: 3
Percent Similarity: 82.836 Percent Identity: 54.478
alignment_block:
US-09-333-159-47 x A26689
Align seg 1/1 to: A26689 from: 1 to: 1378
16 GluMetTrpLeuLeuValAlaValMetPheGlnArgAsnValas 32
51 AAAATGGGTGCTTTTCATGGTGGCAGCTTGTCTATCTGCA...CTGG 97
32 nSerValHis.....MetProThrLysAlaValaspp 43
98 AACTACACATGCTCTTTTGGAAATTCAGCACCACCA.....AACC 138
43 roGluAlaPheMetAsnIleSerGluIleIleGlnHisGlnGlyTrp 59
139 CTGAAGGGAATATGATATAGTCAGATGATTCTCTACTGGGGATACCA 188
60 CysGluGluTrpGluValAlaThrGluAspGlyTrpIleLeuSerValas 76
189 AGTGAATAATATCAAGTTGTGACTGAAGATGGTTATATCTTGAGTTAA 238
76 nArgIleProArgGlyLeuValGlnProLysLysThrGlySerArgProv 93
239 CAGAATTCCTTATGGGAAGAAAAATTCAGGGAACAGAGCCAGAGACCCG 288
93 alValLeuLeuGlnHisGlyLeuValGlyAlaSerAsnTrpIleSer 109
289 TCGTATTTTTCAGCATGGTTTGTCTGCATCAGCTCAAACTGGATCTCC 338
110 AsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAspAlaGlyPheAs 126
339 AACCTGCCCAACAACAGCCTGGCTTTTATCTGSCAGATGCTGGTTATGG 388
126 pValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerArgLysHisLysT 143
389 TGTGTGCTGGGAACAGCAGAGAAATACCTGCTCCAGAAAGAACTTAT 438
143 hrLeuSerIleAspGlnAspGluPheTrpAlaPheSerTrpAspGluMet 159
439 ACTATTCCACAGACTCCGCTTGAATTCCTGGCTTTTACCTTTGATGAATG 488

```

```

301 ACAAACATTGACACAGCAGCTTGGGCTTCTCTCTGGCGATCGTGGCTT 350
125 eAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerArgLysHisL 142
351 TGATGTGTGATGGGAACACAGAGAAACACCTGCTCTCAAGCACACA 400
142 yThrLeuSerIleAspGlnAspGluPheTrpAlaPheSerTrpAspGlu 158
401 AGACTCTTTTCAGTTCTCAGAGATGAGTTCTGGGCTTCAGTTTGTAG 450
159 MetAlaArgPheAspLeuProAlaValIleAsnPheIleLeuGlnLys 175
451 ATGCCAAAATATGACCTACCTGCTTCAATTAACATCATTTTGTATAA 500
175 rGlyGlnGluLysIleTrpValGlyTrpSerGlnGlyThrThrMetG 192
501 TGGCCAAAGAACAAATCTATTATGTGGTCTATTCTCAAGGCTGCACCA 550
192 lyPheIleAlaPheSerThrMetProGluLeuAlaGlnLysIleLysMet 208
551 GTTTCATAGCTTTTTCACAGATGCTGAGTTGGCAAGAGATTAATAAT 600
209 TyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLysSerPro 225
601 TTTTATGCTGTGGCTCCCGTGGTGTCTCTCAATTTGCTTCAGGCCGCT 650
225 yThrLysPheLeuLeuLeuProAspMetMetIleLysGlyLeuPheGly 242
651 ACITCAATTTGGGAGCTTGGCAGACCTCTCTCAAGGACATCTTTGGAC 700
242 ysLysGluPheLeuTrpGlnThrArgPheLeuArgGlnLeuValIleTr 258
701 AGAAGCAATTTCTCCCTCAGAGTCGATGCTGAAGTGGCTGAGCATTC 750
259 LeuCysGlyGlnValIleLeuAspGlnIleCysSerAsnIleMetLeu 275
751 GTTTGGCACCAGCTCATCATGAAGAGCTTGGCAACGCTCTCTTCCT 800
275 uLeuGlyGlyPheAsnThrAsnAsnMetAsnMetSerArgAlaSerVal 292
801 GCTATGTGGCTTCAATCAGAAGAATTTAATATGCTAGATGGATGTGT 850
292 yrAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuHisTr 308
851 ATACACACACTGCTCCTGCGGAATCTCTCGTGAACACATGTTACACTG 900
309 SerGlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpGlySerG 325
901 GGTGAGTCTTTTAAATACCGTAAGCTTCAAGCTTTTGACTGGGGAAG 950
325 uThrLysAsnLeuGluLysCysAsnGlnProThrProValArgTrpArg 342
951 TGAAAAGAAATTTTTCATTACACACAGAGTTTCTCTCCCTCGTACACA 1000
342 alArgAspMetThrValProThrAlaMetTrpThrGlyGlnAspTrp 358
1001 TAAAAAATACGCGTCTACCCACTGCTTGTGGAGCGGGCGGCGACTGG 1050
359 LeuSerAsnProGluAspValLysMetLeuLeuSerGluValThrAsnLe 375
1051 CTGGCAGACATCATGATCATCTATCTTTACTGACTCAAAATCCCAAGTT 1100
375 uIleTrpHisLysAsnIleProGluTrpAlaHisValAspPheIleTrp 392
1101 GGTGTATCACAAGACATTTCTGAGTGGGACCACTCTGGATTTATCTGG 1150
392 lyLeuAspAlaProHisArgMetTrpAsnGluIleIleHisLeuMetGln 408
1151 GTTTGGATGCCCTGTGAACACTGTATCATGAATAATACAGTCTAATGA 1200
409 GlnGluGlu.....ThrAsnLeuSerGlnGly 417

```

[illegible]

```

210 PheAlaLeuAlaProIleAlaThrValIyHisAlaLysSerProGlyTh 226
740 TATCATAGCTCCAGTCGCCACTGTGAAGTATACAAAAGCCCTGTAA 691
226 rLysPheLeuLeuLeuProAspMetMetIleLysGlyLeuPheGlyLysL 243
690 CAACCTTAGTGTATCTCTCAACCATGTTCAAGATTATATTGGTGACA 641
243 ySGLuPheLeuTyGlnThrArgPheLeuArgGlnLeuValIleTyLeu 259
640 AAATATTCTACCCACACAATTTCTTGTATCAATTTCTTGCACCTCAAGTG 591
260 CysGlyGlnValIleLeuAspGlnIleCysSerAsnIleMetLeuLeuLe 276
590 TGTTCCTCGTGAGACACTGAATGTAAATTTGACGAATGCCTTATTATCAT 541
276 uGlyGlyPheAsnThrAsnAsnMetAsnMetSerArgAlaSerValTyRA 293
540 TTGTGGCTTTGACAGCCCAACITTGACATGAGTGGCTTGGATGTGTACG 491
293 laAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuHisTrpSer 309
490 TGTACATATCCGCGAGGACITTCAGTTCAAAACATGCTGCACCTGGACC 441
310 GlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpGlySerGluTh 326
440 CAGCTGTAAATCTGGAATTTTCAAGCTTTTAATTTGGGGAAGTCCAGC 391
326 rLysAsnLeuCluLysCysAsnGlnProThrProValArgTyArgVala 343
390 TCAGATGTAGTGCACTTCAATCAGCCACACCTTCTTCTACTACAAATGTA 341
343 rGAspMetThrValProThrAlaMetTrpThrGlyGlyGlnAspTrpLeu 359
340 CCGCATGAATGTGCCAATGTGAGTGGAGTGGTGGCAATGACTGGTGTG 291
360 SerAsnProGluAspValLysMetLeuLeuSerGluValThrAsnLeuIl 376
290 GCTGACCCCAAGATGTGACCTTTTCTTCTCCAAAACCTTTCTAATCTCAT 241
376 eTyHHisLysAsnIleProGluTrpAlaHisValAspPheIleTrpGlyL 393
240 TTACCAAGAGAGATTTCTCATATATACACTGGGATTTTATCTGGGCAA 191
393 euAspAlaProHisArgMetTyAsnGluIleHisLeuMetGlnGln 409
190 TGAATGCTCTCAAGAAGTTTACAATCAATTAATTTCTATGATGGCAAAA 141
410 GluGlu 411
140 GATAAA 135

```

seq_name: gb_pat1:A01157

```

seq_documentation_block: 1336 bp mRNA PAT 29-JAN-1993
LOCUS A01157
DEFINITION R.norvegicus mRNA for prelingual lipase protein.
ACCESSION A01157
VERSION A01157.1 GI:14748
KEYWORDS lipase.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1336)
AUTHORS Williamson, R
JOURNAL Patent: WO 8500381-A 3 31-JAN-1985;
Celltech Ltd
FEATURES
Location/Qualifiers
1..1336
/organism="Rattus norvegicus"

```

CDS

```

/db_xref="taxon:10116"
/contig="1194"
/codon_start=1
/product="prelingual lipase"
/protein_id="CAA00136.1"
/db_xref="GI:14749"
/db_xref="SWISS-PROT:P04634"
/translation="MWLLITSVITFTGGAGHLFGKLGPNPEANNNISOMITYWGYPCOEYEVVTEDEGYILGYRIPHGKNSINSGKRPVYVYLOHGLIASATNWTNANLPNNSLAFMLADAGYDVLGNSRGNTWSKKNVYSPDSVEFWAFSDENAKYDLPATINFIQVKTGOEKIHYVGHSGOGTTIGFTAFSTNPTLAKKIKEVALAPVATVKYTSPLKISFIPTFLFKLMFGKKMPLPHYTFDDFLGTGECVSRVLDLCSNTLFIKFCGDDKKNLVSRDVIYLGHPAGTSVODFLHWAQLVRSGRFQAFNWGSPQNMLHYNOKTDPEDVDVSAAMTVPVAYWNGNDILADPDQVAMLLPKLSNLLPFKEILAYNHLDFIWMADAPQEVYNEIMISM"
BASE COUNT 409 a 283 c 265 g 379 t
ORIGIN

alignment_scores:
Quality: 1166.50 Length: 399
Ratio: 3.556 Gaps: 1
Percent Similarity: 82.206 Percent Identity: 54.637

alignment_block:
US-09-333-159-47 x A01157
Align seg 1/1 to: A01157 from: 1 to: 1336

16 GluMetTrpLeuLeuIleLeuValAlaTyMetPheGln...ArgAsnVa 31
:::|||||:::|||||:::|||||:::
4 AAGATGGCTGTATTATAACAAGTGTGATATCAACATTCGGAGGTGC 53
:::|||||:::|||||:::|||||:::
31 lAsnSerValHisMetProThrLysAlaValAspProGluAlaPheMetA 48
:::|||||:::|||||:::|||||:::
54 ACATGGCTATTGGAAAACCTGGGTCTCGAACCCTGAACAAATATGA 103
:::|||||:::|||||:::|||||:::
48 snIleSerGluIleIleGlnHisGlnGlyTyProCysGluGluTyGlu 64
:::|||||:::|||||:::|||||:::
104 ATATTAGTCAGATGATAACTTACTTGGGGATATCCATGTCAGAATATGAA 153
:::|||||:::|||||:::|||||:::
65 ValAlaThrGluAspGlyTyIleLeuSerValAsnArgIleProArgG 81
:::|||||:::|||||:::|||||:::
154 GTTGTACTGAAGATGGCTACATTCCTGGGGTCTACAGAAATTCCTCATGG 203
:::|||||:::|||||:::|||||:::
81 yLeuValGlnProLysLysThrGlySerArgProValValLeuLeuGlnH 98
:::|||||:::|||||:::|||||:::
204 GAAGATAATCTCAAAATATAGCAAGACACCTGTGGTGTATTGGCAGC 253
:::|||||:::|||||:::|||||:::
98 lSGlyLeuValGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnAsn 114
:::|||||:::|||||:::|||||:::
254 ATGTTTGATCGCATCAGCCACAACCTGGATTGCAAAATCTACCAACAAC 303
:::|||||:::|||||:::|||||:::
115 SerLeuGlyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGlyAs 131
:::|||||:::|||||:::|||||:::
304 AGCTTGSCCTTTATGTAGCAGACGCTGGCTATGATGTGTGGTGGGAA 353
:::|||||:::|||||:::|||||:::
131 nSerArgGlyAsnAlaTrpSerArgLysHisLysThrLeuSerIleAspG 148
:::|||||:::|||||:::|||||:::
354 CAGTCGGAGAAATACATGCTCCAGGAAAAATGTATACTACTCACCAGACT 403
:::|||||:::|||||:::|||||:::
148 lNAspGluPheTrpAlaPheSerTyAspGluMetAlaArgPheAspLeu 164
:::|||||:::|||||:::|||||:::
404 CAGTTGAATTTCTGGCTTTTCAGCTTTTCAAAATGGCTAAATATGACCTT 453
:::|||||:::|||||:::|||||:::
165 ProAlaValIleAsnPheIleLeuGlnLysThrGlyGlnGluLysIleTy 181
:::|||||:::|||||:::|||||:::
454 CCGGCCCAATAAATAACTTCATTTGACAGAAACCTGGCAAGAGAAGATACA 503
:::|||||:::|||||:::|||||:::
181 rTyrValGlyTySerGlnGlyThrThrMetGlyPheIleAlaPheSerTr 198
:::|||||:::|||||:::|||||:::
504 CTATGTGGTCATCTCAGGGCCACCACCTATGTGTTTCATTCGCTTTCTA 553
:::|||||:::|||||:::|||||:::

```

```

198 hrMetProGluLeuAlaGlnLysIleLysMetTyrPheAlaLeuAlaPro 214
   || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
554 CCATCTACACTGGCCAAATAATACAGACGTTTATGCATAGCTCCA 603
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
215 ILeAlaThrValLysHisAlaLysSerProGlyThrLysPheLeuLeu 231
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
604 GTTGCTACCTGAAGTATACACAAAGTCCCTGCAAAAGATTTCATTAT 653
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
231 uProAspMetMetIleLysGlyLeuPheGlyLysLysGluPheLeuTyrG 248
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
654 TCCTACATTCTTTCAAGCTTATGTTGGCAAGAAATGTTCTCTCCCC 703
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
248 lnThrArgPheLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIle 264
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
704 ACACCTACTTGTAGTACTTCTTGGTACCAAGTGTCTCAGGGAGGT 753
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
265 LeuAspGlnIleCysSerAsnIleMetLeuLeuLeuGlyGlyPheAsnTh 281
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
754 CTAGATCTTCTCTGCACCAACTTTATTCATCTTCTGTGGATTGCACAA 803
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
281 rAsnAsnMetAsnMetSerArgAlaSerValTyrAlaAlaHisThrLeuA 298
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
804 GAAAAACTTAATCTGAGTCTTTGATGTATCTAGGCAATAATCCAG 853
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
298 laGlyThrSerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSer 314
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
854 CAGGGACATCTGTTCAAGACTTCTCCACGTGGCACAGCTGTGTAGATCT 903
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
315 GlyLeuLeuArgAlaPheAspTrpGlySerGluThrLysAsnLeuGluLy 331
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
904 GGGAAATTTCAAGCTTCACTGGGAGGCCATCCAGAACATGTTACA 953
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
331 sCysAsnGlnProThrProValArgTyrArgValArgAspMetThrValP 348
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
954 CTACACCCAGAAAACGCTCTCTGAATATGATGTCTCAGCCATGACTGTGC 1003
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
348 roThrAlaMetTyrThrGlyGlnAspTrpLeuSerAsnProGluAsp 364
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1004 CAGTTGAGTGTGGACGGTGGCAATGACATCTCGCTGATCCCAAGAT 1053
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
365 ValLysMetLeuLeuSerGluValThrAsnLeuIleTyrHisLysAsnII 381
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1054 GTGCCATGCTGCTTCCCAAACTCTCCAACTCTCTGTTCCATAAGGAGAT 1103
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
381 eProGluTrpAlaHisValAspPheIleTrpGlyLeuAspAlaProHisA 398
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1104 TCTTGCCATCAATCACTGACTTCATCTGGCAATGGATGGCCCTCAAG 1153
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
398 rgMetTyrAsnGluIleIleHisLeuMetGlnGlnGluThrAsn 413
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1154 AGGTTTCAATGAGATGATTTCCATGATGGCAAGACTAAAGAAT 1200
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

seq_name: gb_ro2:RNLIP

```

seq_documentation_block:
LOCUS      RNLIP             1355 bp      mRNA           ROD           30-MAR-1995
DEFINITION Rat mRNA for lingual lipase.
ACCESSION X02309
VERSION    X02309.1 GI:56595
KEYWORDS   glycoprotein; lipase.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 1355)
AUTHORS   Docherty,A.J., Bodmer,M.W., Angal,S., Verger,R., Riviere,C.,
            Lowe,P.A., Lyons,A., Emtage,J.S. and Harris,T.J.
TITLE      Molecular cloning and nucleotide sequence of rat lingual lipase
            CDNA
JOURNAL    Nucleic Acids Res. 13 (6), 1891-1903 (1985)
MEDLINE    85215587
COMMENT    Data kindly reviewed (12-FEB-1986) by A.J.P. Docherty.

```

```

FEATURES             Location/Qualifiers
     source            1..1355
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     sig_peptide       49..102
                     /product="put. signal peptide (aa 18 to 1)"
     CDS               49..1236
                     /note="lipase precursor"
                     /codon_start=1
                     /protein_id="CAA26179.1"
                     /db_xref="gi:56596"
                     /db_xref="SWISS-PROT:P04634"
                     /translation="MWLLILTSVSTFGCAHGLFKGLGPNPEANNNISOMITTYGYP
COEYEVTEDCYILGVYRIPHGKKNSENIGKRPVYVLOHGLIASATNMIANLPNNSLA
FMLADAGYDVLGNSRGNTWRSKNVYSPDSVEFWAFSDFEMAKYDLPATINFIYOKT
GOEKIHYVHSGQGTIGTIFAEISTNPILAKKIKTFALAPVATVKVTSPLKKLSIPT
FLKLMFGKMSVDFLPHWQLVRSKGFQFNMGSPSONMLHYNKQTPPEYDVSAMTPV
YLGHNPAQTSVQDFLHWAQLVRSKGFQFNMGSPSONMLHYNKQTPPEYDVSAMTPV
AVWNGNDILADPDQVAMLLPKLSNLLFKELIAYNHLDFIWADAPQEVYNEMISM
AED"
     mat_peptide       103..1233
                     /product="lipase (aa 1-377)"
     misc_feature       505..507
                     /note="aa 153 pot. lipid binding site"
     misc_feature       1355
                     /note="polyA site"
BASE COUNT   400 a 291 c 276 g 387 t 1 others
ORIGIN

```

```

alignment_scores:
  Quality: 1166.50      Length: 399
  Ratio: 3.956         Gaps: 1
  Percent Similarity: 82.206  Percent Identity: 54.637
alignment_block:
US-09-333-159-47 x RNLIP
Align seg 1/1 to: RNLIP from: 1 to: 1355

```

```

16 GluMetTrpLeuLeuIleLeuValAlaTyrMetPheGln...ArgAsnVa 31
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
46 AAGATGTGGCTGTATTATAACAAGTGTGATATCAACATTCGCGAGGTGC 95
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
31 lAsnSerValHisMetProThrLysAlaValAspProGluAlaPheMetA 48
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
96 ACATGCGCTATTTGGAAAACCTGGCTCTGGAACCCCTGAAGCAATATGA 145
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
48 snIleSerGluIleIleGlnHisGlnGlyTyrProCysGluGluTyrGlu 64
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
146 ATATTAGTCAGATGATACTTACTGTGGGATATCCATGTCAAGATATGAA 195
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
65 ValAlaThrGluAspGlyTyrIleLeuSerValAsnArgIleProArgGlu 81
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
196 GTTGTACTCAAGATGCTACATTCCTGGGGTCTACAGAAATTCCTCATGG 245
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
81 yLeuValGlnProLysLysThrGlySerArgProValValLeuLeuGlnH 98
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
246 GAAGAATAATCTCGAAAATATAGCAAGACCTGTGTGTATTGTTCAGC 295
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
98 isGlyLeuValGlyValAserAsnTrpIleSerAsnLeuProAsnAsn 114
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
296 ATGCTTTGATCGCATAGCCACAACTGGATGCAAAATCTACCAACAAAC 345
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
115 SerLeuGlyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGlyAs 131
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
346 AGCTGCGCTTATGCTAGCAGACGCTGGCTATCATGTGTGCTGGGAA 395
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
131 nSerArgGlyAsnAlaTrpSerArgLysHisLysThrLeuSerIleAspG 148
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
396 CAGTCGAGGAATACATGTTCCAGGAAAATGTATCTACTACTCACCAGACT 445
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148 lAspGluPheTrpAlaPheSerTyrAspGluMetAlaArgPheAspLeu 164

```

```

.. |||||
446 CAGTTGAATCTGGCTTTCAGCTTTGATGAATGGCTAAATATGACCT 495
165 ProAlaValIleAsnPhelLeuGlnLysThrGlyGlnGluLysIleTy 181
496 CCCGCCAATAAATCTTATGTCACAGAAACTGGACAGAGACATACA 545
181 rTyrvAlGlyTyrsrGlnGlyThrMetGlyPheIleAlaPheSerT 198
546 CTATGTTGGTCATCTCAGGCCACCACTATTGGTTTCATTGCCCTTCTA 595
198 hrMetProGluLeuAlaGlnLysIleLysMetTyrrPheAlaLeuAlaPro 214
596 CCAATCTACATCTGCGCAAAAATCAAGAGCTTTTATGCAATAGCTCCA 645
215 IleAlaThrValLysHisAlaLysSerProGlyThrLysPheLeuLeu 231
646 GTTGCTACCGTGAAGTATACACAAGTCCCTTGAAAAGATTTCAATAT 695
231 uProAspMetMetIleLysGlyLysPheGlyLysGluPheLeuTyrg 248
696 TCCTACATTTCTTCAAGCTTATGTTGGCAAGAAATGTTCTGCCCC 745
248 lnThrArgPheLeuArgGlnLeuValIleTyrrLeuCysGlyGlnValle 264
746 ACACCTACTTTGATGACTTTCTGGTACCGAAGTGTGCTCACGGGAGTT 795
265 LeuAspGlnIleCysSerAsnIleMetLeuLeuGlyGlyPheAsnTh 281
796 CTAGATCTCTCTGCAGCAACACTTATTCATCTCTGTGGATTTGACA 845
281 rAsnAsnMetAsnMetSerArgAlaSerValTyrrAlaAlaHisThrLeu 298
846 GAAAACCTAAATGTAGTCTGTTGATGTGTATCTAGGGCATAATCCAG 895
298 laGlyThrSerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSer 314
896 CAGGACATCTGTTCAAGACTTCTCCACTGGGCACACAGCTGTTAGTCT 945
315 GlyGluLeuArgAlaPheAspTrpGlySerGluThrLysAsnLeuGlu 331
946 GGGAAATTTCAAGCTTCAACTGGGAAGCCATCCCAAGACATGTTACA 995
331 scYAsnGlnProThrProValArgTyrrArgValArgAspMetThrValp 348
996 CTACAACCAAGAAACGCTCTCTGAATATGATGTCTCAGCCATGACTGTC 1045
348 roThrAlaMetTrpThrGlyGlnAspTrpLeuSerAsnProGluAsp 364
1046 CAGTTGAGTGGAGCGGTGGCAATGACATCTGGCTGATCCCAAGAT 1095
365 ValLysMetLeuLeuSerGluValThrAsnLeuIleTyrrHisLysAsnI 381
1096 GTCCGCAATGCTCTCCCAACTCTCCAACCTCTCTGTTCCATAAGGAGAT 1145
381 eProGluTrpAlaHisValaspPheIleTrpGlyLeuAspAlaProHis 398
1146 TCTTGCTTACCAATCACTGGACTTCATCTGGCAATGGATGCCCTCAAG 1195
398 rgMetTyrrAsnGluIleHisLeuMetGlnGlnGluThrAsn 413
1196 AGGTTTACAAATGAGATGATTTCCATGATGGCAAGAGACATAAAGAT 1242
seq_name: gb_om:CFLIPASE
seq_documentation_block:
LOCUS CFLIPASE 1651 bp mRNA MAN 18-JUN-1998
DEFINITION Canis familiaris mRNA for lipase.
ACCESSION Y13899
VERSION Y13899.1 GI:2204112
KEYWORDS DGL gene; gastric lipase; triacylglycerol lipase.
SOURCE dog.
ORGANISM Canis familiaris

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
Vaganay,S., Joliff,G., Bertaux,O., Toselli,E., Devignes,M.D. and
Benicourt,C.
The complete cDNA sequence encoding dog gastric lipase
DNA Seq. 8 (4), 257-262 (1998)
994:0174
2 (bases 1 to 1651)
Benicourt,C.
Direct Submission
Submitted (10-JUN-1997) C. Benicourt, Ecole Normale Supérieure de
Cachan, L.I.R.B.A (Laboratoire Interdisciplinaire de Recherche en
Biologie Appliquée), 61 avenue du Président Wilson, F-94235 Cachan
Cedex, FRANCE
Location/Qualifiers
1..1651
/organism="Canis familiaris"
/db_xref="taxon:9615"
/dev_stage="adult"
/tissue_type="stomach mucosa"
54..1250
/gene="DGL"
54..1250
/gene="DGL"
/EC_number="3.1.1.3"
/codon_start=1
/product="triacylglycerol lipase"
/protein_id="CAA74198.1"
/db_xref="GI:2204113"
/db_xref="SWISS-PROT:P80035"
/translation="MWLLTAASVISTLTGTHGLFKLHPTNPEVTMNIQSOMITYWGY
PAEYEVVTEGDIILGIDIRIPYGRKNSENIGRRPVAFLOHLLASATNWIISNPNNSL
AFILADAGDVWLGNSRGNTWARRNLYYSPDSVEFWAFSDEMAYDLDPATIDFILKK
TGODKLHYVGHSGOTTIGFIAFSTNPKLAKRIKTEVALAPVATKYTETLLANKMLVP
SEFLKLIIGNKIFYPHHPFDQPLATEVCSRETVDLLCSNALFIICGFDTMNLMSRLD
VILSHNPGATSVQNVLHWSQAVKSRQFADMGSPVONMMHHIHOSSPPYINLTDHVP
IAVWNGNDLLADPHDVLULLSKLPNIYHRKIPYPYNIHLDFIWAMDAPQAVYNEIVSM
MGTDNK"
54..110
/sig_peptide
/gene="DGL"
/evidence="experimental"
111..1247
/mat_peptide
/gene="DGL"
polya_signal 1602..1607
BASE COUNT 432 a 419 c 352 g 448 t
ORIGIN
alignment_scores:
Quality: 1154.00 Length: 398
Ratio: 3.486 Gaps: 5
Percent Similarity: 83.166 Percent Identity: 54.271
alignment_block:
US-09-333-159-47 x CFLIPASE ..
Align seg 1/1 to: CFLIPASE from: 1 to: 1651
17 MetTrpLeuLeuIleLeuValAlaTyrrMetPheGlnArgAsnValAsnSe 33
|||||
54 ATGTGGCTGCTATTAAACAGCGCGGAGTGTGATATCT...ACACTTGGAA 100
|||||
33 rValHis.....MetProThrLysAlaValAspProG 44
|||||
101 CACACATGGTTATTTCGAAAATATACATCCACA.....AACCTG 141
|||||
44 luAlaPheMetAsnIleSerGluIleleGlnHisGlnGlyTyrrProCys 60
|||||
142 AAGTGACCATGAATATAAGTCAGATGATCAGCTACTGGGATACCCAGCT 191
|||||
61 GluGluTyrrGluValAlaThrGluAspGlyTyrrIleLeuSerValAsnAr 77
|||||

```

[illegible]

```

376 eTyrHisLysAsnIleProGluTrpAlaHisValAspPheIleTrpGlyL 393
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
1139 TTACCACAGGAAGATTCCTCTTACATCACTTGGACTTTATCTGGGCCA 1188
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
393 euAspAlaProHisArgMetTyrAsnGluIleIleHisLeuMet 407
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
1189 TGGATGCCCTCAAGCGGTTTACAATGAAATGTTTCCATGATG 1232
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
seq_name: gb_pat1:AR039022

seq_documentation_block:
LOCUS AR039022 1137 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 2 from patent US 5807726.
ACCESSION AR039022
VERSION AR039022.1 GI:5958385
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unknown.
Unclassified.
REFERENCE
1. (bases 1 to 1137)
Blanchard,C., Benicourt,C. and Junien,J.
Nucleic acids encoding dog gastric lipase and their use for the
production of polypeptides
Patent: US 5807726-A 2 15-SEP-1998;
Location/Qualifiers
1. .1137
/organism="unknown"
BASE COUNT 296 a 308 c 249 g 284 t
ORIGIN

alignment_scores:
Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3
Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:
US-09-333-159-47 x AR039022
Align seg 1/1 to: AR039022 from: 1 to: 1137

37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIleI 53
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
19 CCACAC.....AACCTGAAGTGACCATGAATATAAGTCAGATGAT 59
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
53 eGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaThrGluAsp 70
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
60 CACCTACTGGGATACCCAGCTCAGGAATATGAAGTTGTGACCGAAGACG 109
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
70 lTyrIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
110 GTTATATCTCTGGGATCGACAGATTCCTTATGGGAGAAAAATTTCAG 159
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
87 LysThrGlySerArgProValValLeuLeuGlnHisGlyLeuValGlyL 103
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
160 ATATAGGCGCGAGACCTGTGGATTTTTCACACGGTTGTCTCGGATC 209
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
103 yAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleL 120
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
210 AGCCACAACTGGATCTCCAACTGCCCAACACAGCCTGGCCTTCATCC 259
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
120 euAlaSpAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
260 TGGCCACGCCGGGTACAGCGTGTGGCTGGGGAAACAGCAGGGGCAACACC 309
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
310 TGGGCGAGAGGAAATCTGTACTACTCGCCCGACTCCGTCGAATTCGGCG 359
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
153 aPheSerTyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnP 170
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:
360 TTTCAGCTTTGACGAGATGGCTAAATATGACCTTCCCGCCACCATTCACT 409
|||||:|||||:|||||: |||: |||:|||||:|||||:|||||:

```

170 heileLeuGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSer 186
 410 TCATCTTTGAAGAAACGGGACAGCAAGCTACACTAGCTTGGCCATTC 459
 187 GlnGlyThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
 460 CAGGGCACCAACCATGTGTTTCATCGCTTTCCACCACATCCCAAGCTGGC 509
 203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
 510 GAAACGGATCAAAACCTTCTATGATAGCTCCGTTGCCACCGTGAAT 559
 220 IsAlaLysSerProGlyThrLysPheLeuLeuProAspMetMetile 236
 560 ACACGAAACCTGTTAAACAACTCATGCTCGCTCCCTGCTTCTCTTC 609
 237 LysGlyLeuPheGlyLysGluPheLeuTyrGlnThrArgPheLeuAr 253
 610 AAGCTTATATTGGAACAAATATTC...TACCACACCATCTTCTTGA 656
 253 gGln...LeuValIleTyrLysCysGlyGlnValIleLeuAspGlnIleC 269
 657 TCAATTTCTGCCACCGAGGATGCTCCCGGAGCGGTGGATCTCTCT 706
 269 ysSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnAsnMetAsn 285
 707 GCACAAACGCCCTGTTTATCATTTGTGGATTTGACACTATGAATTTGAAC 756
 286 MetSerArgAlaSerValTyrAlaIleHisThrLeuAlaGlyThrSera 302
 757 ATGAGTCGCTGGATGTATCTCTACATCAATCCACGACGACATCGGT 806
 302 lGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgA 319
 807 TCAGACGCTGCTCCACCTGCTCCAGGCTGTTAAGTCTGGGAAGTTCCAAG 856
 319 lApheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
 857 CTTTGTGACTGGGAAGCCCACTGCTCAGACATGATGCTATCATCAGAGC 906
 336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
 907 ATGCTCTCTACTACACCTGACAGACATGCTGTCCTAATCGCAGTGTG 956
 352 pThrGlyGlyClnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
 957 GAACGTTGGCAACGACTGTCTGGCCGACCTCAGCATGTTGACCTTTTGC 1006
 369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAla 385
 1007 TTTCCAGCTCCCAATCTCATTTACACAGGAAGATTCCTCTTACAAT 1056
 386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGl 402
 1057 CACTGTGACTTTATCTGGGCGATGATGCTCCCTCAACCGGTTTACAATGA 1106
 402 uilleIleHisLeuMet 407
 1107 AATTGTTTCCATGATG 1122

seq_name: gb_pat1:AR092633

seq_documentation_block:

LOCUS AR092633 1137 bp DNA

DEFINITION Sequence 2 from patent US 5998189.

ACCESSION AR092633

VERSION AR092633.1 GI:10019385

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1137)

08-SEP-2000

AUTHORS Blanchard,C., Benicourt,C. and Junien,J.
 TITLE Polypeptide derivatives of dog gastric lipase and pharmaceutical
 compositions containing same
 JOURNAL Patent: US 5998189-A 2 07-DEC-1999;
 FEATURES Location/Qualifiers
 source 1..1137
 BASE COUNT 296 a 308 c 249 g 284 t
 ORIGIN

alignment_scores:
 Quality: 1141.50 Length: 372
 Ratio: 3.624 Gaps: 3
 Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:

US-09-333-159-47 x AR092633

Align seg 1/1 to: AR092633 from: 1 to: 1137

37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIleI 53
 19 CCACACA.....AACCTGAAGTGACCATGAATATAAGTCAGATGAT 59
 53 eGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaLalThrGluAspG 70
 60 CACCTACTGGGATACCCAGCTGAGGAATATGAAGTTGTGACCGAAGACG 109
 70 LysThrIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
 110 GTTATATCTTGGGATCGACAGAAATTCCTTATGGGAGGAAAAATTCAGAG 159
 87 LysThrGlySerArgProValValLeuLeuGlnHisGlyLeuValGlyL 103
 160 AATATAGCCGCGAGACCTGTGTGATTTTGCACACACGCTTGTCTGCATC 209
 103 YAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleL 120
 210 AGCCACAACTGGATCTCCAACTGCGCAACACACAGCTGGCTTCATCC 259
 120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
 260 TGCCGACGCGCGGTACGACGTGTGGTGGGAAACACAGCGGGGCAACACC 309
 137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
 310 TGCGCCNGAGGAGATCTGTACTACTCGCCGACTCCGTCGAATCTGGGC 359
 153 apheSerTyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnP 170
 360 TTTTCAGCTTTGACGAGATGGCTAAATATGACCTTCCGCCACCATTTGACT 409
 170 heileLeuGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSer 186
 410 TCATCTTGAAGAAACGGGACAGCAAGCTACACTAGCTTGGCCATTC 459
 187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
 460 CAGGGCACCAACCATGTGTTTCATCGCTTTCCACCACATCCCAAGCTGGC 509
 203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
 510 GAAACGGATCAAAACCTTCTATGATAGCTCCGTTGCCACCGTGAAT 559
 220 IsAlaLysSerProGlyThrLysPheLeuLeuProAspMetMetile 236
 560 ACACGAAACCTGTTAAACAACTCATGCTCGCTCCCTGCTTCTCTTC 609
 237 LysGlyLeuPheGlyLysGluPheLeuTyrGlnThrArgPheLeuAr 253
 610 AAGCTTATATTGGAACAAATATTC...TACCACACCATCTTCTTGA 656

253 gln...LeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleC 269
 657 TCAATTTCTCGCCACCGAGGTATGCTCCGGGAGAGGTGGATCTCCTCT 706
 269 ySerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnMetAsn 285
 707 GCAGCAACGCCCTGTTATCATTTGCTGATTTGACACTATGAAC 756
 286 MetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVa 302
 757 ATGATCTCTGGATGTTATCTGTACATAATCCAGCAGGAACATCGGT 806
 302 IGlAsnIleLeuHisTyrSerGlnAlaValAsnSerGlyGluLeuArgA 319
 807 TCAGAACGTCTCCACTGCTCCAGCGCTTAAGTCTGGGAAGTTCCAAG 856
 319 IaPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
 857 CTTTGTGACTGGGAAGCCAGTTCAGAACATGATGCACATATCATCAGAGC 906
 336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
 907 ATGCTCTCCCTACTACAACTTGACAGACATGATGTGCCAATCGCAGTGTG 956
 352 pThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
 957 GAACGGTGGCAAGACTTGTGCGCCAGCCCTCAGCATGTTGACCTTTTGC 1006
 369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAla 385
 1007 TTTCCAAGCTCCCAATCTCATTTACCACAGGAAGATCTCCTTACAAT 1056
 386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnG1 402
 1057 CACTTGGACTTATCTGGCCATGATGCCCTCAAGCGGTTTACAATGA 1106
 402 uIleIleHisLeuMet 407
 1107 AATTGTTTCCATGATG 1122

seq_name: gb_pat1:AR039023

seq_documentation_block: 1140 bp DNA PAT 29-SEP-1999
 LOCUS AR039023

DEFINITION Sequence 4 from patent US 5807726.

ACCESSION AR039023

VERSION AR039023.1 GI:5958386

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1140)

AUTHORS Blanchard,C., Benicourt,C. and Junien,J.

TITLE Nucleic acids encoding dog gastric lipase and their use for the

production of polypeptides

JOURNAL Patent: US 5807726-A 4 15-SEP-1998;

FEATURES Location/Qualifiers

source 1..1140

BASE COUNT 297 a 308 c 250 g 285 t

ORIGIN

alignment_scores:

Quality: 1141.50 Length: 372

Ratio: 3.624 Gaps: 3

Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:

US-09-333-159-47 x AR039023

Align seg 1/1 to: AR039023 from: 1 to: 1140

37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIleI 53
 22 CCCACA.....AACCTGAAGTGACCATGAATATAAGTCAGATGAT 62
 53 eGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspG 70
 63 CACCTACTGGGATACCCAGCTGAGGAATATCAAGTTGTGACCGAAGAGC 112
 70 LyTyrIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
 113 GTTATATCCTTTGGGATCGACAGATTCCTTATGGGAGGAAAAAATTCAGAG 162
 87 LysThrGlySerArgProValValLeuLeuGlnHisGlyLeuValGlyG1 103
 163 AATATAGCCCGGAGACCTGTTGTCATTTTGGCAACACGGTTTGTCTGCATC 212
 103 yAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleL 120
 213 AGCCACAACCTGGATCTCCAACCTGCCCAACAACAGCCTGGCCTTCATCC 262
 120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
 263 TGGCCGAGCCCGGTACGACGTGTGCTGGGGAACAGCAGGGGCAACACC 312
 137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
 313 TGGGCCAGGAGGAATCTGTACTACTGCGCCGACCTCCGTGCAATTCGGGC 362
 153 aPheSerTyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnP 170
 363 TTTACCTTTGACGAGATGGCTAATATGACCTTCCCGCCACCATTCGACT 412
 170 heIleLeuGlnLysThrGlyGlnGluLysIleTyrValGlyTyrSer 186
 413 TCATCTTGAAGAAAACGGGACAGGACAGCTACACTACGTTGGCCATTCC 462
 187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
 463 CAGGCAACCATTTGGTTTTCATCGCCTTTCCACCAATCCCAAGCTGGC 512
 203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
 513 GAAACGGATCAAAACCTTCTATGCATAGCTCCCGTTGCCACCGTGAAGT 562
 220 IsAlaLysSerProGlyThrLysPheLeuLeuLeuProAspMetMetIle 236
 563 ACACCCAAACCCCTGTTAAACAAACTCATGCTGCTCCTTCCTCTCTTC 612
 237 LysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThrArgPheLeuAr 253
 613 AAGCTTATATTGGAAACAAAATATTC...TACCCACACCACTTCTTTGA 659
 253 gGln...LeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleC 269
 660 TCAATTTCTCGCCACCGAGGTATGCTCCCGGAGAGCGGTGGATCTCCTCT 709
 269 ySerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnMetAsn 285
 710 GCAGCAACGCCCTGTTTATCATTTGTTGATTTGACACTATGAACTTGAAC 759
 286 MetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVa 302
 760 ATGAGTCGCTTGGATGTGTATCTGTACATAATCCAGCAGGAACATCGGT 809
 302 IGlAsnIleLeuHisTyrSerGlnAlaValAsnSerGlyGluLeuArgA 319
 810 TCAGAACGTCTCCACTGCTCCAGGCTGTTAAGTCTGGGAAGTTCCAAG 859
 319 IaPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
 860 CTTTGTACTGGGAAGCCAGTTTCAAGACATGATGCACCTATCATCAGAGC 909
 336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352

263 TGGCCGACGCGGATGACAGCTGTGGTGGGGAACACAGGGGACACACC 312
137 TrpSerArgLysHisLysThrLeuSerLeuAspGlnAspGluPheTpaL 153
313 TGGCCAGGAGGAATCTGTACTACTCCCGGACATCCGTCGAAATTCGGGC 362
153 aPheSerTyTAspGluMetalArgPheAspLeuProAlaValileAsnP 170
363 TTTCAGCTTTGACGAGATGGCTAAATATGACCTTCCCGCCACCATTCGACT 412
170 heileLeuGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSer 186
413 TCATCTTGAAGAAACGGGACAGGACAAGCTACACTACGTTGGCCATTC 462
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
463 CAGGGCACCACCATGGTTTCATCGCTTTTCCACCAATCCCAAGCTGGC 512
203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
513 GAAACGATCAAAACCTCTATGATTAAGTCCCGTGGCACCGTGAAT 562
220 isAlaLysSerProGlyThrLysPheLeuLeuLeuProAspMetMetIle 236
563 ACACCGAACCCTGTTAAACAACATCACTGCTGCTTGGTCTCCTCTTC 612
237 LysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThrArgPheLeuAr 253
613 AAGCTTATTTGGAACAAATATTC...TACCCACACCATCTTTTGA 659
253 gGln...LeuValleTyrLeuCysGlyGlnValleLeuAspGlnIleC 269
660 TCAATTTCTGCCACCAGGATGATCTCCCGGAGACGGTGGATCTCCT 709
269 ysSerAsnIleMeLeuLeuLeuGlyGlyPheAsnThrAsnAspMetAsn 285
710 GCACACGCGCCTGTTATCATTTGTGGATTGACACTATGAACCTGAAC 759
286 MetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVa 302
760 ATGAGTCGCTGGATGTATCTGTCACATAATCCACGAGGAACATCGGT 809
302 lGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgA 319
810 TCAGAAGCTGCTCCACCTGGTCCACAGGCTGTTAAGTCTGGGAAGTTCCA 859
319 laPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
860 CTTTGTGACTGGGAAGCCGAGTTCAGAACATGATGTCACATATCATCACAGC 909
336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
910 ATGCCTCCCTACTACACCTGCACAGACATGCATGTGCCATCGCAGTGTG 959
352 pThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
960 GAACGGTGGCAACGATGCTGTCGGCCAGCCCTCAGCATGTTGACCTTTGC 1009
369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAla 385
1010 TTTCAGAGCTCCCAATCTCATTTACCACAGGAAGATTCCTCTCTACAA 1059
386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGl 402
1060 CACTTGGACTTATCTGGCCATGGATGCCCTCAACGCGTTTACATGA 1109
402 uilleHisLeuMet 407
1110 AATGTTTCCATGATG 1125

OM of: US-09-333-159-47 to: N_Geneseq_0601.* out_format : pfs
 Date: Sep 1, 2001 9:32 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet.p2n.model -DEV=xlpl
 -O=/cgn2_1/USPTO_spool/US09333159/runat_31082001_122331_1419/app_query.fasta_1.488
 -DB=N_Geneseq_0601 -QFMT=fastp -SUFFIX=p2n.rng -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -GAPOP=4.000 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM_ext -MINLEN=0 -MAXLEN=200000000
 -USER=US09333159_ecgn1_1_209 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPPY
 -WAIT -THREADS=1

Search information block:

Query: US-09-333-159-47
 Query length: 423
 Database: N_Geneseq_0601.*
 Database sequences: 730101
 Database length: 313950809
 Search time (sec): 113.380000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAF45132 +		2247.00	4488.65	4.8e-242	1269
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAF45131 +		2247.00	4483.08	9.7e-242	2044
/SID88/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ42310 +		1167.00	2315.42	5.3e-121	1378
/SID88/gcgdata/geneseq/geneseq/NA1985.DAT:AAQ50385 +		1158.50	2298.69	4.6e-120	1336
/SID88/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ68388 +		1141.50	2262.93	4.5e-118	1528
/SID88/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ58915 +		1141.50	2262.93	4.5e-118	1528
/SID88/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ68389 +		1141.50	2262.90	4.5e-118	1531
/SID88/gcgdata/geneseq/geneseq/NA1986.DAT:AAQ60685 +		1124.00	2229.03	3.5e-116	1367
/SID88/gcgdata/geneseq/geneseq/NA1986.DAT:AAQ60566 +		1124.00	2229.03	3.5e-116	1367
/SID88/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ58916 +		1124.00	2229.03	3.5e-116	1367
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ28689 +		695.00	1366.67	3.7e-68	1308
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ28679 +		691.00	1365.95	4.1e-68	699
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ44349 +		596.00	1176.33	1.9e-57	617
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ43458 +		514.50	1000.44	9.4e-48	1718
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ43456 +		502.00	977.01	7.1e-44	1483
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ43444 +		473.50	930.82	7.1e-44	572
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ43459 +		465.00	902.95	2.5e-42	1438
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ43453 +		398.00	774.21	3.7e-35	859
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ43452 +		309.50	598.52	2.3e-25	705
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ60606 +		271.50	473.86	2.0e-18	43676
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ39827 +		230.50	442.55	1.1e-16	549
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ12152 +		173.50	324.94	4.0e-10	707
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ34954 +		160.50	302.63	6.9e-09	509
/SID88/gcgdata/geneseq/geneseq/NA1998.DAT:AAQ35431 +		155.00	302.85	5.9e-09	178
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ43450 +		150.00	276.97	1.9e-07	751
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ43455 +		138.00	264.66	9.0e-07	273
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ43451 +		138.00	254.58	3.3e-06	647
/SID88/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ28743 +		117.00	208.13	0.0013	927
/SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ73350 +		117.00	198.98	0.0041	2028
/SID88/gcgdata/geneseq/geneseq/NA1996.DAT:AAQ28741 +		116.50	207.99	0.0013	861
/SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ73347 +		116.50	207.99	0.0013	861
/SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ74384 +		110.00	173.36	0.1099	5445
/SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ73289 +		105.00	177.07	0.0682	1676
/SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ62362 +		105.00	177.07	0.0682	1676
/SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAQ79322 +		102.00	176.61	0.0724	1041
/SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ29322 +		99.50	131.52	23.50	32042
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAQ30011 +		99.50	131.52	23.50	32042
/SID88/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ89352 +		98.50	168.23	0.2120	1167
/SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ27118 +		98.00	164.97	0.3224	1416
/SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ27117 +		98.00	155.23	1.12	3257

/SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ81263 + 98.00 153.51 1.40
 /SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ69112 + 96.50 147.90 2.88
 /SID88/gcgdata/geneseq/geneseq/NA1998.DAT:AAQ35458 + 95.00 156.98 0.8983
 /SID88/gcgdata/geneseq/geneseq/NA1998.DAT:AAQ35473 + 95.00 152.47 1.60
 /SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AAQ08414 + 95.00 152.19 1.66

seq_name: /SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAF45132

seq_documentation_block:

ID: AAF45132 standard; cDNA; 1269 BP.

XX AAF45132;

DT 30-MAR-2001 (first entry)

DE Human TANGO 294 ORF.

XX Gene therapy: TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety; ss.

OS Homo sapiens.

XX WO200077239-A2.

XX 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159.

XX (MILL-) MILLENNTUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI; 2001-032313/04.

XX P-PSDB; AAB66065.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease.

XX Claim 1; Fig 6; 359pp; English.

XX The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome, disorders,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.

XX Sequence :269 BP; 358 A; 268 C; 309 G; 334 T; 0 other;

alignment_scores:

Quality: 2247.00 Length: 423
 Ratio: 5.312 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-333-159-47 x AAF45132

Align seg 1/1 to: AAF45132 from: 1 to: 1269

```

1 MetLeuGluThrLeuSerArgGlnTrpIleValSerHisArgMetGluMe 17
|||||
1 ATGTTGGAACCTTGTCAAGACAGCTGATGTTCTCACACAGAATGGAAAT 50
|||||
17 tTrpLeuLeuIleLeuValAlaIleTyrMetPheGlnArgAsnValAsnSerV 34
|||||
51 .GTGGCTTCTGATCTGTGGCGGTATATGTTCCAGAGAAATGTGAATTCAG 100
|||||
34 alHisMetProThrLysAlaValAspProGluAlaPheMetAsnIleSer 50
|||||
101 TACATATGCCAACTTAAGCTGTGACCCAGAGCAATTCATGAATATTAGT 150
|||||
51 GluIleLeuGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaIle 67
|||||
151 GAAATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCAAC 200
|||||
67 rGluAspGlyTyrIleLeuSerValAsnArgIleProArgGlyLeuValG 84
|||||
201 TGAAGATGGGTATATCCTTCTGTAAACAGGATTCCCTCGAGGCCCTAGTGC 250
|||||
84 InProLysLysThrGlySerArgProValValLeuLeuGlnHisGlyLeu 100
|||||
251 AACCTAGACAGACAGGTTCCAGGCCTGTGGTTACTGCAGCATGGCCTA 300
|||||
101 ValGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuG1 117
|||||
301 GTTGGAGGTGTAGCAACTGGATTCTCAACCTGCCCAACAATAGCCTGGG 350
|||||
117 yPheIleLeuAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgG 134
|||||
351 CTTTCATCTGGCAGATGCTGTTTGGCTGTGGATGGGGAACAGCAGGG 400
|||||
134 lyAsnAlaTrpSerArgLysHisLysThrLeuSerIleAspGlnAspGlu 150
|||||
401 GAAACGCTGTCTCGAAACACACAGACACTCTCCATAGACCAAGATGAG 450
|||||
151 PheTrpAlaPheSerTyrAspGluMetAlaArgPheAspLeuProAlaIa 167
|||||
451 TTCCTGGGCTTTCAGTTATGATGAGATGGCTAGGTTTGACCTTCCTGCAGT 500
|||||
167 lIleAsnPheIleLeuGlnLysThrGlyGlnGluLysIleTyrTyrValG 184
|||||
501 GATAAATCTTATTTTGCAGAAACCGGCCAGGAAAGATCTATTATGTGCG 550
|||||
184 lyTyrSerGlnGlyThrThrMetGlyPheIleAlaPheSerThrMetPro 200
|||||
551 GCTATTTCACAGGGCACCACCATGGCTTTATTGCAATTTCCACCATGCCA 600
|||||
201 GluLeuAlaGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaIle 217
|||||
601 GAGCTGGCTCAGAAATCAAAATGATTTTGTGTTAGCACCCATAGCCAC 650
|||||
217 rValLysHisAlaLysSerProGlyThrLysPheLeuLeuProAspM 234
|||||
651 TGTTAAGCATCAAAAGCCCGGACCAAAATTTTGTGTCGCAGATA 700
|||||
234 etMetIleLysGlyLeuPheGlyLysGlyPheLeuTyrGlnThrArg 250
|||||
701 TGATGATCAAGGGATTGTTGGCAAAAGAAATTTCTGTATCAGACCCAGA 750
|||||
251 PheLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspG1 267
|||||
751 TTTCTCAGACAACTGTGTTTACCTTGTGGCCAGGTGATCTTGTATCA 800
|||||
267 nIleCysSerAsnIleMetLeuLeuLeuGlyPheAsnThrAsnAsnM 284
|||||
801 GATTGTAGTAAATATGTTACTTCTGGGTGGATTCAACACCACCAATA 850
|||||
284 etAsnMetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThr 300
|||||
851 TGAACATGAGCGGAGCAAGTGTATGCTGCCACACACTCTTGCTGGAACA 900
|||||

```

```

301 SerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLe 317
|||||
901 TCTGTGCAAAATATTCTACACTGGAGCCAGGCAGCTGAATTCGTGTAAC 950
|||||
317 uArgAlaPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnG 334
|||||
951 CCGGGCATTTTACTGGGGAGTGAGACCAAAATCTGGAAAATGCAATC 1000
|||||
334 InProThrProValArgTyrArgValArgAspMetThrValProThrAla 350
|||||
1001 AGCCAACCTCTGTAAAGGTACAGAGTCAGAGATATGACGGTCCCTACAGCA 1050
|||||
351 MetTrpThrGlyGlnAspTrpLeuSerAsnProGluAspValLysMe 367
|||||
1051 ATGTGGACAGAGGTCAGGACTGGCTTTCAAATCCAGAAGACGTGAAAAT 1100
|||||
367 tLeuLeuSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluT 384
|||||
1101 GCTGCTCTCTGAGGTGACCAACCTCATCTACCATAGAATATTCCTGAAT 1150
|||||
384 rPalaHisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyr 400
|||||
1151 GGGCTCACGTGGATTTCATCTGGGGTTTGGATGCTCTCCACCGTATGTAC 1200
|||||
401 AsnGluIleIleHisLeuMetGlnGlnGluThrAsnLeuSerGlnG1 417
|||||
1201 AATGAATCATCCATCTGATGCAGCAGGAGGAGACCAACCTTTCCAGGG 1250
|||||
417 yArgCysGluAlaValLeu 423
|||||
1251 ACGGTGTGAGGCCGTATTG 1269
|||||

```

seq_name: /SIDSR/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF45131

seq_documentation_block:

ID AAF45131 standard; cDNA; 2044 BP.

XX AAF45131;

XX 30-MAR-2001 (first entry)

XX Human TANGO 294 cDNA.

XX Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety; ss.

OS Homo sapiens.

XX WO200077239-A2.

XX 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159.

XX (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI: 2001-032313/04.

XX DR P-PSDB; AAB66065.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease.

XX Claim 1; Fig 6; 359pp; English.

xx The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 xx

Sequence 2044 BP; 618 A; 401 C; 460 G; 565 T; 0 other;

alignment_scores:
 Quality: 2247.00 Length: 423
 Ratio: 5.312 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-333-159-47 x AAF45131 ..

Align seg 1/1 to: AAF45131 from: 1 to: 2044

1 MetLeuGluThrLeuSerArgGlnTrpIleValSerHisArgMetGluMe 17
 126 ATGTTGGAAACCTTGCAAGACAGTGGATTGCTCACACAGAATGGAAT 175
 17 trpIleuLeuIleuValAlaTyrMetPheGlnArgAsnValAsnServ 34
 176 GTGGCTTCGATTCTGGTGGCGGTATATGTTCCAGAGAAATGTGAATTCAG 225
 34 aHisMetProThrLysAlaValAspProGluAlaPheMetAsnIleSer 50
 226 TACATATGCCAACTAAAGCTGTGGACCCAGAACGATTCATGAATATTAGT 275
 51 GluIleIleGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaLath 67
 276 GAAATCATCCAACTCAAGGCTATCCCTGTGAGGAATATGAATCGCAAC 325
 67 rGluaspGlyTyrIleLeuSerValAsnArgIleProArgGlyLeuValG 84
 326 TGAAGATGGGTATATCTCTTCTGTTAACAGGATTCCTCGAGGCTTAGTGC 375
 84 InProLysLysThrGlySerArgProValValLeuLeuGlnHisGlyLeu 100
 376 AACCTAAGANGACAGGTCCAGGCTGTGGTGTACTGCAGCATGGCCTTA 425
 101 ValGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuG1 117
 426 GTTGGAGGTGTAGCAACTGGATTTCCAACTGCCCAACATAGCTGGG 475
 117 yPheIleLeuAlaaspAlaGlyPheaspValTrpMetGlyAsnSerArgG 134
 476 CTTTCATTCTGGCAGATGCTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 525
 134 lYAsnAlaTrpSerArgLysHisLysThrLeuSerIleaspGlnaspGlu 150
 526 GAAACGCTTGGTCTCGAAACACAGACACTCTCCATAGACCAAGATGAG 575
 151 PheTrpAlaPheSerTyrAspGluMetAlaArgPheAspLeuProAlaVa 167
 576 TTCTGGGCTTTTCAGTTATGATGAGATGGCTAGGTTTGACCTTCCTGCAGT 625
 167 lIleAsnPheIleLeuGlnLysThrGlyGlnGluLysIleTyrTyrValG 184
 626 GATAAAGCTTTATTTGCGAGAAACGGGCGAGGAGAAAGATCTATTATGTCG 675
 184 lYTySerGlnGlyThrThrMetGlyPheIleAlaPheSerThrMetPro 200
 |||||||||||||||||||||||||||||||||||||||||||||||||||

676 GCTATTTCACAGGGCACCACCATGGGCTTTATTGTCATTTTCCACCATGCCA 725
 201 GluLeuAlaGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaTh 217
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 726 GAGCTGCTCAGAAATCAAAATGATTATTGTTAGCACCATGCCAC 775
 217 rVallyHisAlaLysSerProGlyThrLysPheLeuLeuLeuProAspM 234
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 776 TGTAAAGCATGCAAAAGCCCGGAGCCAAATTTTGTGTGCGCCAGATA 825
 234 etMetIleLysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThrArg 250
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 826 TGATCAACAAGGATTGTTGGCAAAAGAATTTCTGTATCAGACCAGA 875
 251 PheLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspG1 267
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 876 TTTCTCAGACAACTTGTTATTATTTACCTTTGTGGCCAGGTGATTCTTATCA 925
 267 nIleCysSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnAsm 284
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 926 GATTGTAGTAATATCATGTACTTCTGGTGGGATTCAACACCAACAATA 975
 284 etAsnMetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThr 300
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 976 TGAACATGAGCGCAGCAAGTGTATATGCTGCCACACACTCTTCTGCGAACA 1025
 301 SerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLe 317
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 1026 TCTGTGCAAAATATTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAAC 1075
 317 uArgAlaPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnG 334
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 1076 CGGGCATTTGACTGGGGAGTGAGACCAAAAATCTGGAATAATGCAATC 1125
 334 InProThrProValArgTyrArgValArgAspMetThrValProThrAla 350
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 1126 AGCCAACCTCCTGTAAAGTACAGATGACAGATATGACGGTCCCTCAGCA 1175
 351 MetTrpThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMe 367
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 1176 ATGTGACAGAGGAGTCCAGGCTGGCTTTTCAATCCAGAGACGTGAAAT 1225
 367 tLeuLeuSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluT 384
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 1226 GCTGCTCTCTGAGGTGACCAACCTCATCTACCATAGAATATTCCTGAAT 1275
 384 tpAlaHisValaspPheIleTyrGlyLeuaspAlaProHisArgMetTyr 400
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 1276 GGGCTCAGCTGGATTTCATCTGGGGTGTGGATGCTCCTCACCGTATGTAC 1325
 401 AsnGluIleIleHisLeuMetGlnGlnGluThrAsnLeuSerGlnG1 417
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 1326 RATGAATCATCATCTGATGCGAGCAGGAGGAGACCAACCTTTCCAGGG 1375
 417 yArgCysGluAlaValLeu 423
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 1376 ACGGTCTGAGGCGGTATTG 1394

seq_name: /SID8/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ42310

seq_documentation_block:

ID AAQ42310 standard; DNA; 1378 BP..

XX

AC AAQ42310;

XX

DT 20-SEP-1993 (first entry)

XX

DE RGL precursor.

XX

KW Rabbit gastric lipase; RGL; pRGLN2.1; fat; bioconversion;

XX hydrolysis; transesterification; ds.

XX OS Oryctolagus cuniculus.

16-DEC-1992; 92FR-0015201.
(LJOU) INST RECH JOUVEINAL.
Benicourt C, Blanchard C, Junien J;
WPI: 1994-217890/26.
P-PSDB; AAR56871.
Recombinant canine gastric lipase and nucleic acid encoding it -
are used for improving absorption of ingested fat, treating
mucoviscidosis etc. and in enzymatic bio-conversions
Claim 2; Fig 8; 52pp; French.
The sequence given below is the sequence of figure 8, altered
according to the amendments described on page 2 of the appended
letter.
CGL is used to improve absorption of ingested fat, in healthy and
sick patients (e.g. having altered levels of gastric lipase); to
treat conditions associated with insufficiency (or lack) of lipases,
esp. mucoviscidosis or exocrine pancreatic insufficiency and partic.
where immobilised, for bioconversions, e.g. hydrolysis or
transesterification (other mammalian gastric lipases, or derivs.,
can be used in this application).
Sequence 1531 BP: 395 A; 386 C; 329 G; 421 T; 0 other;

```

alignment_scores:
    Quality: 1141.50      Length: 372
    Ratio: 3.024          Gaps: 3
    Percent Similarity: 56.452

alignment_block:
    US-09-333-159-47 x AAQ60389 ..

Align seg 1/1 to: AAQ60389 from: 1 to: 1531

37  ProThrLysAlaValaspProGluAlaPheMetasnIleSerGluIleI 53
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
22  CCACACA.....AACCTGAAGTCACCATGAATATAAGTCAGATGAT 62
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
53  eGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaThrGluAsp 70
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
63  CACCTACTGGGGATACCCACCTCAGGAATATGAAGTTCTGACCGAAGACG 112
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
70  LYTyrIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
113  GTTATATCCTTGGGATCGCAGAGAAATTCCTATGGGAGGAAAAATTCAGAG 162
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
87  LysThrLysSerArgProValValLeuLeuGlnHisGlyLeuValGlyG 103
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
163  AATATAGCGGGAGACCTGTTGCATTTTGGCAACACGGTTTGTCTGCATC 212
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
103  yAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIle 120
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
213  AGCCACAACTGGATCTCCAACTCTGCCCAACACAGCCTGGCCTTCATCC 262
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
120  euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
263  TGGCCGACGCGGGTACGACGTGTGGCTGGGGAAACAGCAGGGGGCAACACC 312
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
137  TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
313  TGGGCCACGAGGAATCTGTACTACTCCGCCGACCTCCGTCGAATTCCTGGC 362
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
153  aPheSerTyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnP 170
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
363  TTTTCAGTTTGACGAGATGGCTAAATATGACCTTCCCGCCACCATGTACT 412
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|
170  baIleLeuGlnLysThrGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 186
||||| :|||||:| :|||||:| :|||||:| :|||||:| :|||||:|

```



```
|||||
323 TGGATTTCACAGCTGCCAACAACAGCTTGCTTCATTCTGGCAGATGC 372
123 aGlyPheAspValTTPMetGlyAsnSerArgGlyAsnAlaTrpSerArgL 140
|||||
373 TGGTTATGATGTGGCTGGCGAAGCAGAGAGAAACACCTGGCGCAGAA 422
140 ySHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSerTyr 156
423 GAACTTGTTACTATTCCACCAAGATTGATTGCTGGCTTCAGCTTT 472
157 AspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLeuG1 173
473 GATGAATGGCTAAATATGACCTCCAGCCCAATCGACTTCATTGTAAA 522
173 nLysThrGlyGlnGluLysIleTyrValGlyTyrSerGlnGlyThrT 190
523 GAAACTGGACAGAGCAGCTACACTATGTTGGCCATTGCCAGGGCACCA 572
190 hrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLysIle 206
573 CCATTGGTTTATTGCTTTCCACCAATCCAGCCTGGCTAAAGAATC 622
207 LysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLysSe 223
623 AAAACCTTCATGCTCTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 672
223 rProGlyThrLysPheLeuLeuLeuProAspMetMetIleLysGlyLeuP 240
673 CTTTATAACAACCTAGATTGTTGCTCAATCCCTCTTCAGATTATAT 722
240 hGlyLysLysGluPheLeuTyrGlnThrArgPheLeuArgGlnLeuVal 256
723 TTGGTGACAAATAATTCTACCCACACAACCTCTTTGATCAATTTCTGCT 772
257 IleTyrLeuGlyGlnValIleLeuAspGlnIleCysSerAsnIleMe 273
773 ACTGAAGTGTGCTCCCGTAGATGCTGAATCCCTTTGCAGCAATGCCCT 822
273 tLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetAsnMetSerArgAlas 290
823 ATTTATAATTTGGATTGACAGTAAAGACTTTAACACGAGTCCCTTGG 872
290 erValTyrAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeu 306
873 ATGTGTATCTATCATAATCCAGCAGGAACCTCTGTTCAAAACATGTC 922
307 HisTrpSerGlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpG1 323
923 CAITGGACCCAGCGTGTAAAGTCTGGAAATCCCAAGCTTATGACTGGGG 972
323 ySerGluThrLysAsnLeuGluLysCysAsnGlnProThrProValArgT 340
973 AAGCCAGTTTCAGATAGGATGACATATGATCAGTCCCAACCTCCCTACT 1022
340 yArgValArgAspMetThrValProThrAlaMetTrpThrGlyGln 356
1023 ACAATGTGACACCCATGAATGTACCAATTTGCAGTGTGGAACGGTGGCAAG 1072
357 AspTrpLeuSerAsnProGluAspValLysMetLeuLeuSerGluValTh 373
1073 GACCTGTGGCTGACCCCAAGATGTTGGCTTTTGTCTCCAAACCTCCC 1122
373 rAsnLeuIleTyrHisLysAsnIleProGluTrpAlaHisValAspPheI 390
1123 CAATCTTATTTACCACAGGAGATTCTTTTACAACTACCTTGGACTTTA 1172
390 leTrpGlyLeuAspAlaProHisArgMetTyrAsnGluIleIleHisLeu 406
1173 TCTGGCAATGGATGCCCTCAAGAAGTTTACAATGACATGTTCTTCTATG 1222
407 MetGlnGlnGlu 411
|||||
```

```
1223 ATATCAAGATAAA 1237
seq_name: /SIDS8/gcgdata/geneseq/geneseq/NA1986.DAT:AA60566
seq_documentation_block:
ID   AA60566 standard; DNA; 1367 BP.
XX
AC   AA60566;
XX
DT   22-AUG-1991 (first entry)
XX
DE   Sequence encoding human pregastric lipase.
XX
KW   Cystic fibrosis therapy; enzyme; lipase deficiency; ss.
XX
OS   Homo sapiens.
XX
FH   Key Location/Qualifiers
FT   CDS 47..103
FT     /*tag= a
FT     mat_peptide 104..1243
FT     /*tag= b
XX
PN   WO8601532-A.
XX
PD   13-MAR-1986.
XX
PF   15-AUG-1985; 85WO-GB00364.
XX
PR   21-AUG-1984; 84GB-0021210.
PR   15-AUG-1985; 85WO-GB00364.
PR   01-JAN-1986; 86GB-0008897.
XX
PA   (CELL-) CELTECH LTD.
PA   (LOWE/) LOWE P A.
XX
PI   Lowe PA;
XX
DR   WPI; 1986-081634/12.
DR   P-PSDB; AAP60658.
XX
PT   New gastric lipase protein, esp. of human origin - for treating
PT   lipase deficiency, and DNA sequences coding for it
XX
PS   Disclosure; Fig 3; 39pp; English.
XX
CC   The inventors claim a pregastric lipase protein and a gene encoding
CC   it. Gastric lipase protein is useful for oral administration to
CC   treat lipase deficiency, e.g. cystic fibrosis or pancreatitis.
XX
SQ   Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;

alignment_scores:
Quality: 1124.00 Length: 405
Ratio: 3.416 Gaps: 2
Percent Similarity: 81.235 Percent Identity: 52.099

alignment_block:
US-09-333-159-47 x AA60566 ..
Align seg 1/1 to: AA60566 from: 1 to: 1367

14 ArgMetGluMetTrpLeuLeuIleLeuValAlaTyrMetPheGlnArgAs 30
||| ..|||
38 AGTCCAAATGTGGCTCTTTTACAAATGGCAAGTTTGATA..... 79
30 nValAsnSerValHisMetProThrLysAlaVal..... 41
|||
80 .....TCTGTACTGGGGACTACACATGGTTTGTGGAAATTCATC 122
|||
42 .....AspProGluAlaPheMetAsnIleSerGluIleIleGlnHisGln 56
|||||
```

```

123 CTGGAAGCCCTGAGTACTATGAACATTAGTCAGATGATTACTATTATGG 172
57 GlyTyrProCysGluGluTyrGluValAlaThrGluAspGlyTyrIleLe 73
173 GGATACCAATGAAGAATATGAAGTTGTGACTGAAGATGGTTATATCT 222
73 userValAsnArgIleProArgGlyLeuValGlnProLysLysThrGlyS 90
223 TGAAGTCAATAGATTCTTATGGGAAGAAAAATTCAGGGAATACAGGCC 272
90 erArgProValValLeuLeuGlnHisGlyLeuValGlyGlyAlaSerAsn 106
273 AGAGACCTGTTGTGTTTGGCAGTGTGTTGCTTCATCAGCCCAAAAC 322
107 TrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAspAl 123
323 TGGATTTCACCTCCGGAACAACAGCCTTCCTTCATTCTGGCAGATGC 372
123 aglyPheAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerArgL 140
373 TGGTTATGATGTGGCTGGCGACAGCAGAGGAAACACCTGGGCCAGAA 422
140 ysHisLysThrLeuSerLeuAspGlnAspGluPheTrpAlaPheSerTyr 156
423 GAAACTTGTACTATTACACAGATTCCAGTTGAATTCCTGGCTTTCAGCTT 472
157 AspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLeuG 173
473 GATGAATGGCTTAATATGACCTTCACGCCACAAATCGACTTCATGTAAA 522
173 nLysThrGlyGlnGluLysIleTyrValGlyTyrSerGlnGlyThr 190
523 GAAAACTGGACAGACAGAGCTACACTATGTTGGCCATCCCGAGGCCA 572
190 hrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLysIle 206
573 CCATTGGTTTATTCCTTTCCACCAATCCAGCCTGGCTGAAAGAAATC 622
207 LysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLysSe 223
623 AAAACCTTCTATGCTCTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 672
223 rProGlyThrLysPheLeuLeuProAspMetMetIleLysGlyLeuP 240
673 CCTTATAACAAATAGATTGTTCTCAATCCCTCTCAAGTTTATAT 722
240 heGlyLysGluPheLeuTyrGlnThrArgPheLeuArgGlnLeuVal 256
723 TTGGTGACAAATATCTTACCACACACTCTTTGATCAATTTCTGCT 772
257 IleTyrLeuCysGlyGlnValIleLeuAspGlnIleCysSerAsnIleMe 273
773 ACTGAAGTGTGCTCCGTGAGATGCTGAATCTCCTTTGCAGCAATGCCT 822
273 tLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetSerArgAlaAs 290
823 ATTTATAATTTGGATTGGAGTTGACAGTAAGACTTTAACACAGCTGCT 872
290 erValTyrAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeu 306
873 ATGTGTATCATCATATCCAGCAGGAACTCTCTGTTCAAAACATGTT 922
307 HisTrpSerGlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpG 323
923 CATTTGACCCAGGCTGTAAAGTCTGGAAATTCGAAGCTTATGACTGGG 972
323 ySerGluThrLysAsnLeuGluLysCysAsnGlnProThrProValArgT 340
973 AGCCCAATGAGATGAGTACATGATGATGATGATGATGATGATGATGAT 1022
340 yrArgValArgAspMetThrValProThrAlaMetTrpThrGlyGlyGln 356
1023 ACAATGTACACCCATGAATGATACCAATGTCAGTGTGGAACGGTGGCA 1072

```

```

357 AspTrpLeuSerAspProGluAspValLysMetLeuLeuSerGluValTh 373
1073 GACCTGTTGGCTGACCCCAAGATGTTGGCTTTCCTCCAAACTCCC 1122
373 rAsnLeuIleTyrHisLysAsnIleProGluTrpAlaHisValAspPheI 390
1123 CAATCTTATTTACCACAAGGAGATTCTTTTACAATCACTTGGACTTTA 1172
390 leTrpGlyLeuAspAlaProHisArgMetTyrAsnGluIleIleHisLeu 406
1173 TCTGGCCAATGGATGCCCTCAAGAAGTTTACAATGACATTGTTCTATG 1222
407 MetGlnGlnGluGlu 411
1223 ATATCAGAAGATAAA 1237

```

seq_name: /SIDS8/gcgdata/geneseq/geneseq/NA1996.DAT:AAT58916

seq_documentation_block:

ID AAT58916 standard; DNA; 1367 BP.

XX AAT58916;

XX AC

XX XX

XX 19-AUG-1997 (first entry)

XX Human gastric lipase coding sequence.

XX DE

XX Duodenal; gastric; lipase; transgenic; plant; recombinant; extract; ss;

XX KW food; absorption; fat; pancreatic; cystic fibrosis; exocrine; dairy;

XX KW hydrolysis; trans-esterification; substrate; enzyme; biofuel; human.

XX OS Homo sapiens.

XX XX

XX Key Location/Qualifiers

XX CDS 47..1243

XX FT /*tag= a

XX FT /product= gastric lipase

XX FT sig_peptide 47..103

XX FT /*tag= b

XX FT mat_peptide 104..1240

XX FT /*tag= c

XX XX

XX PN W09633277-A2.

XX XX

XX PD 24-OCT-1996.

XX XX

XX PF 19-APR-1996; 96WO-FR00606.

XX XX

XX PR 20-APR-1995; 95FR-0004754.

XX XX

XX PA (BIOC-) BIOCEM SA.

XX PA (LJOU) INST RECH JOUVEINAL.

XX XX

XX PI Baudino S, Benicourt C, Cudrey C, Gruber V, Lenee P;

XX PI Merot B;

XX XX

XX DR WPI; 1996-485783/48.

XX DR P-PSDB; AAW09383.

XX XX

XX PT Recombinant human or canine pre-duodenal lipase prodn. in transgenic

XX PT plants - useful for facilitating absorption of fat, as bio-catalysts

XX PT and for prodn. of bio-fuel

XX PS Claim 3; Fig 4; 130pp; French.

XX XX

XX CC This is the nucleotide sequence encoding the human pre-duodenal (i.e.

XX CC gastric) lipase enzyme. The sequence can be used to generate transgenic

XX CC plants producing recombinant lipase in an enzymatically active form.

XX CC Alternatively bases encoding amino acids 20-23 or 20-73 (1-4 or 1-54 of

XX CC the mature protein, respectively) can be deleted to form the derivatives

XX CC designated delta-4 or delta-54 respectively. Plants, or their extracts,

XX CC expressing the lipases or the truncated derivatives, can be used:

CC (a) as pharmaceuticals or food to facilitate absorption of fat, either
 CC in healthy subjects or in patients with inadequate levels of
 CC gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine
 CC pancreatic insufficiency, in the elderly or in patients undergoing medical
 CC treatment which alters fat absorption; (b) for performing industrial or
 CC agricultural reactions, e.g. in processing of fats or in the dairy
 CC industry, for hydrolysis or transesterification reactions, etc., where
 CC the plant material may provide both enzyme and substrate. The transgenic
 CC plants can also be used for biofuel production.
 xx
 SQ Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;

alignment_scores:
 Quality: 1124.00 Length: 405
 Ratio: 3.416 Gaps: 2
 Percent Similarity: 81.235 Percent Identity: 52.099

alignment_block:

US-09-333-159-47 x AAT58916 ..

Align seg 1/1 to: AAT58916 from: 1 to: 1367

14 ArgMetGluMetTrpLeuLeuLeuValAlaValMetPheGlnArgAs 30
 38 AGGTCCAAATGTGGCTGCTTTTAAACAATGGCAAGTTTGATA..... 79
 30 nValAsnSerValHisMetProThrLysAlaVal..... 41
 80TCGTACTGGGGACTACATGCTTGTGTTGGAAATATACATC 122
 42AspProGluAlaPheMetAsnLeuSerGluIleGlnHisGln 56
 123 CTGGAAGCCCTGAAGTCACTAATGAACATTAGTCAGATGATTACTATTGG 172
 57 GlyTyrProCysGluGluTyrGluValAlaValAlaThrGluAspGlyTyrIleLe 73
 173 GGATACCAATGAAGAAATATGAAGTTGTGACTGAAGATGTTATATCT 222
 73 uSerValAsnArgIleProArgGlyLeuValGlnProLysLysThrGlyS 90
 223 TGAAGTCAATAGATTCTTATGGGAGAGAAATTCAGGGNATACAGGCC 272
 90 eArgProValValLeuGlnHisGlyLeuValGlyAlaSerAsn 106
 273 AGAGACCTGTGTGTTTGGCAGCATGGTTGCTTGTCATCAGCCACAAAC 322
 107 TrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAspAl 123
 323 TGGATTTCCAACTCCCGCAACACAGCCTTGCCTTCATTCTGGCAGATGC 372
 123 aGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerArgL 140
 373 TGGTATGATGCTGTGGTGGCCACACAGCAGAGGAAACACCTGGGCCAGAA 422
 140 ysHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSerTyr 156
 423 GAAACTGTCTACTATTACACAGATTCACTGTAATTCCTGGCTTTCAGCTTT 472
 157 AspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLeuGlu 173
 473 GATGAATGGCTTAATATGACTTCCAGCCACAAATCGACTTCATGTGAAA 522
 173 nLysThrGlyGlnGluLysIleTyrValGlyTyrSerGlnGlyThr 190
 523 GAAACTGGACAGAGCAGCTACACTATGTGGCCATTCACAGGGCACCA 572
 190 hrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLysIle 206
 573 CCATTTGGTTTATTGCTCTTTCCACCAATCCAGCTGGCTTAAAGAATC 622
 207 LysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLysSe 223
 ||||||||.....

623 AAAACCTTCTATGCTCTAGCTCTGTTGCCACTGTGAAGTATACAAAAG 672
 223 rProGlyThrLysPheLeuLeuLeuProAspMetMetIleLysGlyLeup 240
 673 CCTATATAACAACTTAGATTGTCTCAATCCCTCTCAAGTTATAT 722
 240 heGlyLysLysGluPheLeuTyrGlnThrArgPheLeuArgGlnLeuVal 256
 723 TTGGTGCAAAATATTCTACCCACACAACTTCTTTGATCAATTTCTTGTCT 772
 257 IleTyrLeuCysGlyGlnValIleLeuAspGlnIleCysSerAsnIle 273
 773 ACTGAAGTGTGCTCCCTGGAGATGCTCAATCTCTTTGCAGCAATGCC 822
 273 tLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetSerArgAlas 290
 823 ATTATATTTGTGGATTGTGACAGTAAAGACTTTAACACGAGTCGCTTGG 872
 290 erValTyrAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeu 306
 873 ATGTGTATCTATCATATAATCCAGCAGGAACCTCTGTTCAAAACATGTC 922
 307 HisTrpSerGlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpGlu 323
 923 CATTTGGACCCAGGCTGTAAAGTCTGGGAAATTCAGCTTATGACTGGGG 972
 323 ySerGluThrLysAsnLeuGluLysCysAsnGlnProThrProValArgT 340
 973 AGCCCAAGTTCAGATAGGATGCACTATGATCAGTCCCAACCTCCCTACT 1022
 340 yArgValAlaArgAspMetThrValProThrAlaMetTrpThrGlyGlyGln 356
 1023 ACAATGTGACAGCCATGAATGTACCAATTCAGTGTGGAACGGTGGCAAG 1072
 357 AspTrpLeuSerAsnProGluAspValLysMetLeuLeuSerGluValTh 373
 1073 GACCTGTGGCTGACCCCAAGATGTGGCCTTTTGTCTCCAAACTCCC 1122
 373 rAsnLeuIleTyrHisLysAsnIleProGluTrpAlaHisValAspPheI 390
 1123 CAATCTTATTACCAAGAGATTCCTTTTACAACTCAGCTGGACTTAA 1172
 390 leTrpGlyLeuAspAlaProHisArgMetTyrAsnGluIleIleHisLeu 406
 1173 TCTGGCAATGGATGCCCTCAAGAAGTTTACAATGACATTTGTTCTATG 1222
 407 MetGlnGlnGluGlu 411
 1223 ATATCAGAGATAAA 1237
 seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF28689
 seq_documentation_block:
 ID: AAF28689 standard; cDNA; 1308 BP.
 XX
 AC: AAF28689;
 XX
 DT: 05-APR-2001 (first entry)
 XX
 DE: Human protein HP03372 coding sequence #2.
 XX
 KW: Human; hydrophobic domain; immune deficiency; autoimmune disorder;
 KW: allergy; tissue growth; regeneration; wound healing; burn; tumour;
 KW: periodontal disease; thrombolytic condition; haemostatic condition;
 KW: infection; ss.
 XX
 OS: Homo sapiens.
 XX
 PN: WO200102563-A2.
 XX
 PD: 11-JAN-2001.
 XX
 PF: 16-JUN-2000; 2000WO-JP03943.

CC antibodies. They may be useful for treatment of autoimmune disorders
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
 CC osteoporosis, osteoarthritis, central nervous system disorders
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
 CC disease), tumours, bacterial, fungal or viral infections, depression and
 CC priortiasis. AAA45926 to AAA45931 represent linker variants which are given
 CC in the exemplification of the present invention.
 XX
 SQ Sequence 617 BP; 181 A; 125 C; 139 G; 172 T; 0 other;

alignment_scores:
 Quality: 596.00 Length: 190
 Ratio: 3.772 Gaps: 2
 Percent Similarity: 83.158 Percent Identity: 60.526

alignment_block:

US-09-333-159-47 x AAA44349 ..

Align seg 1/1 to: AAA44349 from: 1 to: 617

```

16 GluMetTrpLeuLeuValAlaTyrMetPheGlnArgAsnValas 32
   :::::::::::::::::::::
54 AAATGTGGCTGCTTTTAACTAGGCAAGTTGATA..... 89
   :::::::::::::::::::::
32 nSerValHisMetProThrLysAlaVal.....A 42
   :::::::::::::::::::::
90 .TCTCTACTGGGGACTACACATGTTGTTGGAAAAATTACATCTCGAA 138
   :::::::::::::::::::::
42 sProGluAlaPheMetAsnIleSerGluIleIleGlnHisGlnGlyTyr 58
   :::::::::::::::::::::
139 GCCCTGGAAGTACTAGTAACTATGTCAGATGATTACTATTGGGGATAC 188
   :::::::::::::::::::::
59 ProCysGluGluTyrGluValAlaThrGluAspGlyTyrIleLeuSerVa 75
   :::::::::::::::::::::
189 CCAATGAAGATATGAAGTGTGACTGAAGTGTATATCTTGAAGT 238
   :::::::::::::::::::::
75 IasnArgIleProArgGlyLeuValGlnProLysLysThrGlySerArgp 92
   :::::::::::::::::::::
239 CAATAGAATTCCTTATGGAGAGAAAATTCAGGGAATACAGGCCAGAGAC 288
   :::::::::::::::::::::
92 roValValLeuLeuGlnHisGlyLeuValGlyGlyAlaSerAsnTrpIle 108
   :::::::::::::::::::::
289 CPGTTGTGTTTTCAGCATGCTGTTGCTGCATCAGCCAACTGGGATT 338
   :::::::::::::::::::::
109 SerAsnLeuProAsnAspSerLeuGlyPheIleLeuAlaAspAlaGlyph 125
   :::::::::::::::::::::
339 TCCAACTGCGGAGAACACAGCTTGCCTTCATCTGGCAGATGCTGGTTA 388
   :::::::::::::::::::::
125 easpValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerArgLysHisL 142
   :::::::::::::::::::::
389 TGATGTGTGGTGGCAACAGCAGAGAGAAACACCTGGGCCAGAAACT 438
   :::::::::::::::::::::
142 ysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSerTyrAspGlu 158
   :::::::::::::::::::::
439 TGTACTATTCACCAATTCAGTTGAATTCGGGCTTTCAGCTTTCATGAA 488
   :::::::::::::::::::::
159 MetAlaArgPheAspLeuProAlaValIleAsnPheIleLeuGlnLysTh 175
   :::::::::::::::::::::
489 ATGGCTAAATATGACCTCCAGCCACATCGACTTCATGTAAAGAAAC 538
   :::::::::::::::::::::
175 rGlyGlnGluLysIleTyrTyrValGlyTyrSerGlnGlyThrMetG 192
   :::::::::::::::::::::
539 TGGACAGAGAGCAGCTACACTATGTGGCCATCCAGGGCCACCATTCG 588
   :::::::::::::::::::::
192 lyPheIleAlaPheSerThr 198
   :::::::::::::::::::::
589 GTTTATTGGCTTTTCCACC 608
  
```

seq_name: /SIDS8/gcgdata/geneseq/geneseq/NA2000.DAT:AA234958

seq_documentation_block:
 ID AA234958 standard; cDNA; 1718 BP.

XX AA234958;
 AC
 XX 28-FEB-2000 (first entry)
 DT
 XX
 DE Soybean acid triacylglycerol lipase cDNA.
 XX
 KW Triacylglycerol lipase; soybean; fatty acid; seed oil;
 KW vegetable oil; transgenic plant; ss.
 XX
 OS Glycine max.

XX
 FH Key Location/Qualifiers
 FT CDS 184..1416
 FT /*tag= a

XX WO9555883-A2.

XX 04-NOV-1999.

XX 29-APR-1999; 99WO-US09280.

XX 30-APR-1999; 98US-0083688.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Cahoon EB, Cahoon RE, Kinney AJ, Rafalski JA;

XX WPI: 2000-062036/05.

XX P-PSDB; AAY32309.

XX Novel plant triacylglycerol lipase polynucleotides used to alter the
 PT level of the enzyme in transgenic plants

XX Claim 2; Page 48; 65pp; English.

CC This is the nucleotide sequence of a contig assembled from the
 CC entire cDNA insert in clone ssl.pk0022.a1 and a portion of the
 CC cDNA insert in clone sdp3c.pk004.n3 encoding the entire soybean
 CC acid triacylglycerol lipase (TAGL) (see AAY32309). The clones were
 CC isolated from soybean developing pod and seedling (5-10 day post
 CC germination) cDNA libraries. Novel acid and neutral TAGL polypeptides
 CC (see AAY32301-17) and polynucleotides (see AA234950-66) from corn,
 CC Catalpa, rice, soybean and wheat tissues are disclosed. The enzymes
 CC may be prepared recombinantly and used to raise antibodies, which
 CC are used for detecting the enzymes in situ in cells or in vitro in
 CC cell extracts. The polynucleotides may be used to create transgenic
 CC plants in which the TAGL levels are present at higher or lower levels
 CC than normal, or in cell types or developmental processes where they are
 CC not normally found. This would alter the level of triacylglycerol and
 CC cholesterol esters found in those cells. Accumulation of fatty acids
 CC with unusual structures may be a positive phenotype in plants used for
 CC foods. In addition, it may be desirable to eliminate expression of TAGL
 CC genes for certain applications. TAGL enzymes may also be useful for the
 CC processing of plant seed oils and for the development of novel seed
 CC oils. The TAGL enzymes can also be used as targets to facilitate the
 CC design and/or identification of inhibitors of those enzymes that may be
 CC useful as herbicides. This is desirable because inhibition of the
 CC activity of either of the enzymes could lead to an inhibition of plant
 CC growth. The polynucleotides also serve as a source of probes and
 CC primers, which are useful for genetic mapping, as markers for traits
 CC linked to those genes, and to isolate homologous sequences from other
 CC species.

XX SQ Sequence 1718 BP; 492 A; 353 C; 344 G; 529 T; 0 other;

alignment_scores:

Quality: 514.50 Length: 429
 Ratio: 1.913 Gaps: 17
 Percent Similarity: 62.704 Percent Identity: 32.634

alignment_block:

US-09-333-159-47 x AA234958

Align seg 1/1 to: AA234958 from: 1 to: 1718

```

1 MetLeuGlnThrLeuSerArgGlnTrpIleValSerHisArgMetGluMe 17
232 GTCTCAACAAGTGTGCTCGTCAGGACACGCTTCAAGCGGTGGCAACTT 281
17 ttrpLeuIleLeuValAlaTyrMetPheGlnArgAsnValAsnSerV 34
282 A.....GGCAGAAACATCAAC..... 297
34 alHisMetProThrLysAlaValAspProGluAlaPhe...MetAsnIle 49
298 .....CCTTCAGTGTATGGCATATGTGCC 321
50 SerGluIleIleGlnHisGlnGlyTyrProCysGluGluTyrGluValAl 66
322 TCTTCTCTCATTTGGCAT...GGATACAAGTGTCAAGAACACGAGTTAC 368
66 aThrGluAspGlyTyrIleLeuSerValAsnArgIleProArgGlyLeuV 83
369 AACTGATGATGGTTACATCTCGACCTGCGCAAGGATCCCGAGAGGTCGAG 418
83 alGlnProLysLysThrGlySerArg.....ProValValLeuLeuGln 97
419 GTAAAGCAGTGGGAGTGGGACAGGAAGCAACACGAGTGT...ATACAA 465
98 HisGlyLeuValGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnAs 114
466 CATGGAGTCTTGTAGATGTATGACATGCGCTTCTAAACCCACGAGCA 515
114 nSerLeuGlyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGlyA 131
516 AGATCTCCGCTGTGATTTAGCTGATATGATGATTTGACGTGTGGATTGCAA 565
131 snSerArgGlyAsnAlaTrpSerArgLysHisLysThrLeuSerIleAsp 147
566 ACACAGAGGAACACAGATATAGTCGCGACACATCTCATTTGGACCCCTCT 615
148 GlnAspGluPheTrpAlaPheSerTyrAspGluMetAlaArgPheAspLe 164
616 AGCCAGCCCTATTGGAATTTGCTTGGGATGAACTTGTCTCCTATGATTT 665
164 uProAlaValIleAsnPheIleLeuGlnLysThrGlyGlnGluLysIleT 181
666 CCCTGGCGTGTAAATTAATGATGTTACGCCAAACGGGCGAG...AAGATCA 712
181 yTyrValGlyTyrSerGlnGlyThrThrMetGlyPheIleAlaPheSer 197
713 ATTACGTTGGCCATTTCATTGGGAACCTTTGGTAGCTTTGGCATCCTCTCG 762
198 ThrMetProGluLeuAlaGlnLysIleLysMetTyrPheAlaLeuAlaPr 214
763 GAA...GGAAATTTGGTTACCCAGCTGAATTCAGCAGCCCTTTGTAGCCCC 809
214 oileAlaThrValLysHisAlaLysSerProGlyThrLysPheLeuLeuL 231
810 TATAGCCTATTTAAGCCACATGAATACA.....GCACCTTGGTGTCTTGTG 853
231 euProAspMetMet.....IleLysGlyLeuPheGlyLysLysGlu 244
854 CACCCAAAGTCTTTGTTGGTGAGATCACTACCCTCTTCGCTAGCAGAA 903
245 Phe.....LeuTyrGlnThrArgPheLeuArgGlnLeuValII 257
904 TTTAATCCAAAGGGTGTGATGTTGATGCTTCTCAAGTCT..... 945
257 eTyrLeuCysGlyGlnValIleLeuAspGlnIleCysSerAsnIleMetL 274
946 ....CTCTGTGCTCACCCCTGGGATAGAC.....TGCTATGACTTGTGTA 985

```

```

274 euLeuLeuGlyGlyPheAsnThrAsnAsnMetAsnMetSerArgAlaSer 290
986 CTGCACCTAACTGTAATAATGTC...TGCCTCAATCTTCAACTGTTGAT 1032
291 ValTyrAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuH1 307
1033 CTATTCCTGATGATGAGCCTCAGTCAGTCACATCAACAAGAACATGTTGCA 1082
307 sTrpSerGlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpGlyS 324
1083 CTGTGGCTCAGACGTTAGACTTGGGGGTTGACAAATTCATATATGTGA 1132
324 erGluThrLysAsnLeuGluLysCysAsnGlnProThrProValArgTyr 340
1133 GACCAGACTATAACATTATGACACTATGAGAAATATTTCTCCAATCTAT 1182
341 ArgValArgAspMet.....ThrValProThrAlaMetTrpThrGlyG1 355
1183 AACCTTTCACACATCCCCACGATCTCCCTCTCTTCATTAGCTATGGTGG 1232
355 yGlnAspTrpLeuSerAsnProGluAspValLysMetLeuLeuSerGluV 372
1233 AAGAGATGCACTTTCAGATGTCGCGTATGTTGAGAATTTGCTTGATAAA. 1281
372 alThrAsnLeuIleTyrHisLysAsn.....Ile 381
1282 .....CTCAAGTTCATGATGAGAAACAGCGCAGCGTTCAGTTTCATC 1323
382 ProGluTrpAlaHisValAspPheIleTrpGlyLeuAspAlaProHisAr 398
1324 CAGGAATATGCTCATGCTGACTACATATGGGTTCATGCAATGCCAAGGACTT 1373
398 gMetTyrAsnGluIleIleHisLeuMetGlnGlnGlu 410
1374 GGTGTATATGCTGCTTCTTTCATTTTTCATCATCAAA 1410

```

seq_name: /SIDSE/gcgdata/geneseq/geneseqn/NA2000.DAT:AA234956

seq_documentation_block:

ID AA234956 standard; cdna; 1483 BP.

XX	AA234956;	
XX		
DT	28-FEB-2000 (first entry)	
XX		
DE	Rice acid triacylglycerol lipase cDNA.	
XX		
KW	Triacylglycerol lipase; rice; fatty acid; seed oil;	
KW	vegetable oil; transgenic plant; ss.	
XX		
OS	Oryza sativa.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	32..1264
FT		/*tag- a
XX		
PN	WO9955883-A2.	
XX		
PD	04-NOV-1995.	
XX		
PF	29-APR-1995;	99WO-US09280.
XX		
PR	30-APR-1995;	98US-0083688.
XX		
PA	(DUPO) DU PONT DE NEMOURS & CO E I.	
XX		
PI	Cahoon EB, Cahoon RE, Kinney AJ, Rafalski JA;	
XX		
DR	WPI; 2000-062036/05.	
DR	P-PSDB; AAY32307.	
XX		
PT	Novel plant triacylglycerol lipase polynucleotides used to alter the	

OM of: US-09-333-159-47 to: Issued_Patents_NA:* out_format : pfs
 Date: Sep 1, 2001 8:18 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEI=framet.p2n.model -DEV=xlp
 -Q/cgn2_1/USPTO_spool/US09333159/runat_31082001_122330_1399/app_query.fasta_1.488
 -DB=issued_Patents_NA -OFMT=fastap -SUFFIX=p2n.rni
 -GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000
 -LOOPT=0.000 -GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000
 -XGAPEXT=0.500 -GAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000
 -YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1
 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCLALIGN=200
 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
 -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=200000000
 -USER=US09333159_cgn1_1_62 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:

Query: US-09-333-159-47
 Query length: 423
 Database: Issued_Patents_NA:*
 Database sequences: 324599
 Database length: 94655562
 Search time (sec): 69.100000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-227-108-2	1	1141.50	2461.52	1137	1.7e-129
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-073-674-2	1	1141.50	2461.52	1137	1.7e-129
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-227-108-4	1	1141.50	2461.49	1140	1.7e-129
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-073-674-4	1	1141.50	2461.49	1140	1.7e-129
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-227-108-6	1	1141.50	2461.42	1146	1.7e-129
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-073-674-6	1	1141.50	2461.42	1146	1.7e-129
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-227-108-1	1	1141.50	2457.71	1152	2.8e-129
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-073-674-1	1	1141.50	2457.71	1152	2.8e-129
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-356-952-12	1	271.50	510.09	43676	8.5e-21
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-751-782-2	1	155.00	326.10	178	1.5e-10
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-073-674-2	1	155.00	326.10	178	1.5e-10
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-484-108-21	1	105.00	187.71	1676	0.0077
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-073-674-21	1	105.00	187.71	1676	0.0077
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-602-359A-24	1	102.00	187.29	1041	0.0081
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-073-674-24	1	98.50	178.16	1167	0.0262
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-232-519-1	1	98.50	178.16	1167	0.0262
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-115-446-1	1	92.00	153.99	2521	0.5820
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-115-446-5	1	92.00	153.96	2526	0.5839
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-471-119A-1	1	92.00	116.26	2160	73.51
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-080-386-3	1	91.00	153.79	2160	0.5967
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-742-185-102	1	91.00	153.79	2160	0.5967
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-968-543C-1	1	89.00	136.17	6027	5.72
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-755-559-2	1	85.00	140.54	2180	3.27
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-210-474-2	1	85.00	140.54	2180	3.27
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-602-359A-30	1	84.50	149.29	1017	1.06
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-759-945-1	1	84.50	137.05	2625	5.11
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-07-707-367-1	1	84.50	133.78	3383	7.78
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-386-607-1	1	84.00	138.47	2160	4.26
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-07-882-329-4	1	82.50	141.46	1329	2.90
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-183-213-4	1	82.50	141.46	1329	2.90
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-446-803-4	1	82.50	140.29	1455	3.37
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-861-837-4	1	82.50	140.29	1455	3.37
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-600-656-4	1	82.50	140.29	1455	3.37
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-170-670-9	1	82.50	140.29	1455	3.37
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-09-170-670-14	1	82.50	140.29	1455	3.37
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-193-068-9	1	82.50	140.29	1455	3.37
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-193-068-13	1	82.50	140.29	1455	3.37
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-09-183-412-9	1	82.50	140.29	1455	3.37
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-183-412-13	1	82.50	140.29	1455	3.37

/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-500-908A-11 + 82.50 140.26 3.39
 /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-683-838A-11 + 82.50 140.26 3.39
 /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-483-101-1 + 82.50 122.45 33.26
 /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-348-891A-1 + 82.50 109.43 176.52
 /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-905-817-1 + 82.50 109.43 176.52
 /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-227-108-2

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-227-108-2

seq_documentation_block:
 ; Sequence 2, Application US/08227108
 ; Patent No. 5807726
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/227,108
 ; FILING DATE: 03-APR-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fanucci, Allan A.
 ; REGISTRATION NUMBER: 30,256
 ; REFERENCE/DOCKET NUMBER: 7620-033
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 790-9090
 ; TELEFAX: 212 869-8864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1137 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: Double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; NAME/KEY: CDS
 ; LOCATION: 1..1137
 ; US-08-227-108-2

alignment_scores:
 Quality: 1141.50 Length: 372
 Ratio: 3.624 Gaps: 3
 Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:
 US-09-333-159-47 x US-08-227-108-2

Align seg 1/1 to: US-08-227-108-2 from: 1 to: 1137

37 ProThrLysAlaValaspProGluAlaPheMetAsnIleSerGluIle1 53
 |||||
 19 CCCACA.....AACCCCTGAGTGACCATGAATTAAGTCAGATGAT 59
 |||||
 53 eGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspG 70
 |||||
 60 CACCTACTGGGATACCCAGCTGAGGAATATGAAGTTGTGACCGAAGACG 109
 |||||
 70 lYTyrlleuSerValAsnArgIleProArgGlyLeuValGlnProLys 86

```
|||||..... |||
110 GTTATATCTTGGGATCGACAGATTCTTATGGGAGGAAATTCAGAG 159
87 LysThrGlySerArgProValValLeuLeuGlnHisGlyLeuValGlyG1 103
|||||..... |||
160 AATATAGCGCGGACGACCTGTGCTGCTTTGCAACACGCTTGTGCTGCATC 209
103 yAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleL 120
|||||..... |||
210 AGCCACAACTGGATCTCCAACTGCGCCCAACACAGCCTGGCCTTCATCC 259
120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
|||||..... |||
260 TGGCGGACGCGGTACGACGTGTGGCTGGGACACACGACGCGGACACACC 309
137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
|||||..... |||
310 TGGCCAGGAGGAACTGTACTACTCGCCGCACTCCGTCGAATTCGTGGC 359
153 aPheSerTyTAspGluMetAlaArgPheAspLeuProAlaValIleAsn 170
|||||..... |||
360 TTCAGCTTTGACGAGATGCTAATATGACCTTCCGCCCACTTGCATGACT 409
170 heIleLeuGlnLysThrGlyGlnGluLysIleTyTrValGlyTyTrSer 186
|||||..... |||
410 TCATCTTTGAAGAAACGCGACGACGACAACTACACTACGTGGCCATTC 459
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
|||||..... |||
460 CAGGCGACCACTTGTGTTTCATCGCTTTTCCACCAATCCCAAGCTGCG 509
203 aGlnLysIleLysMetTyTrPheAlaLeuAlaProIleAlaThrValIsh 220
|||||..... |||
510 GAACGAGTCAAACTCTATGCTAGCTCCGTTGCCACCGTGAAGT 559
220 IsAlaLysSerProGlyThrLysPheLeuLeuProAspMetMetIle 236
|||||..... |||
560 ACACCGAAACCTGTTAAACAACTCATGCTCGTCCCTTCCTTCCTTC 609
237 LysGlyLeuPheGlyLysLysGluPheLeuTyTrGlnThrArgPheLeuAr 253
|||||..... |||
610 AAGCTTATATTTGGAACAAATATTC...TACCACACCACTCTTTGA 656
253 gGln...LeuValIleTyTrLeuCysGlyGlnValIleLeuAspGlnIleC 269
|||||..... |||
657 TCATTTCTCGCCACCGAGTGTCTCCCGGAGCGGTGGATCTCCTCT 706
269 ySerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetAsn 285
|||||..... |||
707 GCAGCAACGCCCTGTTATCATTTGTGGATTTGACACTATGAACCTTCAAC 756
286 MetSerArgAlaSerValTyTrAlaAlaHisThrLeuAlaGlyThrSera 302
|||||..... |||
757 ATGAGTCGCTTGGATGTGTATCTCTCACAATAATCCAGCAACATCGGT 806
302 lGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgA 319
|||||..... |||
807 TCAGACAGTCTCCACTGGTCCCGAGGTGTTAAGTCTGGGAAGTCTCAAG 856
319 laPheAspTrpGlySerGluTyTrLysAsnLeuGluLysCysAsnGlnPro 335
|||||..... |||
857 CTTTGTAGTGGGAAGCCGATTCAGAACATGATGCATCATCATCAGAGC 906
336 ThrProValArgTyTrArgValArgAspMetThrValProThrAlaMetTr 352
|||||..... |||
907 ATGCTCTCTCTACTACACCTGACAGACATGCATGTGCCAATCGCAGTGT 956
352 pThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
|||||..... |||
957 GAACGCTGGCAACGACTTGTGCGCCACCTCAGCATGTGTGACCTTTGCG 1006
369 euSerGluValThrAsnLeuIleTyTrHisLysAsnIleProGluTrpAla 385
|||||..... |||
```

```
1007 TTTCACAGCTCCCAATCTCATTTACCACAGGAAGATTCTCTCTTACAA 1056
386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyTrAsnG1 402
|||||..... |||
1057 CACTTGCATTTATCTGGGCGATGATGCTCAAGCGGTTTACAATGA 1106
402 uIleIleHisLeuMet 407
|||||..... |||
1107 AATGTTTCCATGATG 1122
seq_name: /cgn_6/ptodata/1/ina/5B_COMB.seq:US-09-073-674-2
seq_documentation_block:
; Sequence 2, Application US/09073674
; Patent No. 5958189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-DI-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1137 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1137
; US-09-073-674-2
alignment_scores:
Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3
Percent Similarity: 84.677 Percent Identity: 56.452
alignment_block:
US-09-333-159-47 x US-09-073-674-2
Align seq 1/1 to: US-09-073-674-2 from: 1 to: 1137
37 ProThriysalaValAspProGluAlaPheMetAsnIleSerGluIle1 53
|||||..... |||
19 CCCACA.....AACCTGAAGTGACCATGATATAGTCAGATCAT 59
53 eGlnHisGlnGlyTyTrProCysGluGluTyTrGluValAlaThrGluAspG 70
```

```
109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000
```

```
957 GAACGGTGGCAACGACTGCTGGCGGACCCTCAGCATGTTGACCTTTGTC 1006
369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAla 385
1007 TTTCCAGCTCCCAATCTCATTTACCAACAGAGATCTCTCTTACAAT 1056
386 HisValaspheleTyrGlyLeuaspAlaProHisArgMetTyrAsnG1 402
1057 CACTGGAGCTTTATCTGGGCGCATGGATCCCTCAAGCGGTTTACAATGA 1106
402 uileIleHisLeuMet 407
1107 AATGTTTCCATGATG 1122
seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-227-108-4
seq_documentation_block:
; Sequence 4, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1140 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1140
; US-08-227-108-4
alignment_scores:
Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3
Percent Similarity: 84.677 Percent Identity: 56.452
alignment_block:
US-09-333-159-47 x US-08-227-108-4
Align seg 1/1 to: US-08-227-108-4 from: 1 to: 1140
37 ProThrLysAlaValaspheleTyrGlyLeuaspAlaProHisArgMetAsnIleSerGluIle 53
```


Align seg 1/1 to: US-09-073-674-4 from: 1 to: 1140

```
37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIle1 53
   |||||
   :|||
   :|||
22 CCCACA.....AACCTGAAGTGACCATGAATATAAGTCAGATGAT 62
   |||
53 eelnHISGlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspG 70
   |||
63 CACCTACTGGGGATACCCAGCTGGAGATATGAAGTTGTGACCGAAGACG 112
   |||
70 lYtYrIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
   |||
113 GTTATATCTTGGGATCAGAGAAATTCCTTATGGGAGGAAATTCAGAG 162
   |||
87 LysThrGlySerArgProValValLeuLeuGlnHISGlyLeuValGlyG 103
   |||
163 AATATAGCCGGAGACCTGTTGCTATTTTGCACACGCTTGTGTCGCATC 212
   |||
103 ValaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleL 120
   |||
213 ACCCAACAACCTGGATCTCCACCTCCCAACACAGCCTGGCCTTCATCC 262
   |||
120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
   |||
263 TGGCCGACGCGGGTACGACGTGTGGCTGGGAAACAGCAGGGGCAACACC 312
   |||
137 TrpSerArgLysHISLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
   |||
313 TGGGCGAGGAGAACTCTACTACTCTGCGCGGACTCCGTCGAATTCGGGC 362
   |||
153 aPheSerTrpAspGluMetAlaArgPheAspLeuProAlaValIleAsn 170
   |||
363 TTTCAGCTTTGACGAGATGGCTAAATATGACCTTCCGCGCACATGACT 412
   |||
170 heIleLeuGlnLysThrGlyGlnGluLysIleTyrValGlyTyrSer 186
   |||
413 TCATCTTGAAGAAACGGGACGAGGACAGCTACACTAGCTTGGCCATCC 462
   |||
187 GlnGlyThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
   |||
463 CAGGCGACCAACATTTGGTTTTCATCGCTTTCCACCAATCCCAAGCTGC 512
   |||
203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
   |||
513 GAACGGATCAAAACCTTCTATGATAGCTCCGTTGCCACCGTGAAGT 562
   |||
220 lAlaLysSerProGlyThrLysPheLeuLeuProAspMetMetIle 236
   |||
563 ACACCGAACCCTGTTAAACAACATCATCTCGTCCCTCGTTCTCTTC 612
   |||
237 LysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThrArgPheLeuAr 253
   |||
613 AAGCTATATTTGGAACAAAATATTC...TACCACACACCACTCTTTGA 659
   |||
253 gLIn...LeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleC 269
   |||
660 TCAATTTCTGCGCCAGGATAGCTCCCGGAGAGCGGTGGATCTCCTCT 709
   |||
269 ySerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetAsn 285
   |||
710 GCAGCAACCCCTGTTATCATTTGCGATTGACACTATGAACITGAAC 759
   |||
286 MetSerArgAlaSerValTyrAlaAlaHISThrLeuAlaGlyThrSerVa 302
   |||
760 ATGAGTCGCTTGGATGTGTATCTGTACATAATCCAGCAGGAACATCGGT 809
   |||
302 lGlnAsnIleLeuHISTrpSerGlnAlaValaValaAsnSerGlyGluLeuArgA 319
   |||
810 TCAGAACGTCGTCCTAGCTGTCGAGGCTTAAAGTCTGGGAAGTCCCAAG 859
   |||
319 lAaPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
   |||
```

```
860 CTTTGTACTGGGAGCCAGTTCAGAACATGATGCACATCATCATCAGAGC 909
336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
   |||
910 ATGCTCTCCTACTACAACTGACAGACATGTCATGTGCCAATCGCAGTGTG 959
352 pThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
   |||
960 GAACGGTGGCAACGACTTGTGGCCGACCTCCAGATGTGACCTTTTGC 1009
369 euSerGluValThrAsnLeuIleTyrHISLysAsnIleProGluTrpAla 385
   |||
1010 TTTCCAAGCTCCCAATCTCATTTACCAGGAAGATTCCTCTTACAAT 1059
386 HISValAspPheIleTrpGlyLeuAspAlaProHISArgMetTyrAsnG 402
   |||
1060 CACTTGGACATTTATCTGGCCATGATGCCCTCAAGCGGTTTACAATGA 1109
402 uIleIleHISLeuMet 407
   |||
1110 AATGTTTCCATGATG 1125
seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-227-108-6
seq_documentation_block:
: Sequence 6, Application US/08227108
: Patent No. 5807726
: GENERAL INFORMATION:
: APPLICANT: Blanchard, Claire
: APPLICANT: Benicourt, Claude
: APPLICANT: Junien, Jean-Louis
: TITLE OF INVENTION: Recombinant Dog Gastric Lipase
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/227.108
: FILING DATE: 03-APR-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Fanucci, Allan A.
: REGISTRATION NUMBER: 30,256
: REFERENCE/DOCKET NUMBER: 7620-033
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212 790-9090
: TELEFAX: 212 869-8864/9741
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1146 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
US-08-227-108-6
```

alignment_scores:

Quality:	1141.50	Length:	372
Ratio:	3.624	Gaps:	3
Percent Similarity:	84.677	Percent Identity:	56.452

alignment_block:

US-09-333-159-47 x US-08-227-108-6

Align seg 1/1 to: US-08-227-108-6 from: 1 to: 1146

```
37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIleil 53
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
19 CCCACA.....AACCTTGAGTGACCATGATATATAGTCAGATGAT 59
53 eGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspG 70
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
60 CACCTACTGGGATACCCAGCTGAGGAATATGAAGTTGTGACCGAAGACG 109
70 LysThrIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
110 GTTATATCTTGGGATCGACAGAAATCTTATGGGAGGAAATTCAGAG 159
87 LysThrGlySerArgProValValLeuLeuGlnHisGlyLeuValGlyG 103
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
160 AATATAGCCGAGACCTGTTGCAATTTTGGCAACAGGTTGCTGCGCATC 209
103 yAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleL 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
210 AGCCACAACTGGATCTCCACCTGCCCAACACAGCGCTGCGCTTCATCC 259
120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
260 TGGCCGACGCGGTTACAGCTGTGGCTGGGGAAACAGCAGGGGCAACACC 309
137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
310 TGGCCGACGAGGATCTGTACTACTCGCCGACCTCCGTCGGAATCTGGGC 359
153 aPheSerTyrAspGluMetalaArgPheAspLeuProAlaValIleAsnP 170
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
360 TTTCAGCTTTCAGGAGATGGCTAAATATGACCTTCCGCCACCATGACT 409
170 heIleLeuGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSer 186
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
410 TCATCTTGAAGAAACGGGACGAGCAGACAGTACACTAGTTGGCCATTC 459
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
460 CAGGCGACCCACCATGTTTCATCGCTTTCACCAATCCACAGCTGGC 509
203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
510 GAAACGGATCAAAACCTTCTATGATAGCTCCGTTGCCACCGTGAAGT 559
220 IsAlaLysSerProGlyThrLysPheLeuLeuLeuProAspMetMetile 236
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
560 ACACCGAAACCCCTGTTAAACAACTCATGCTCGTCCCTTCTCTCTTC 609
237 LysGlyLeuPheGlyLysGluPheLeuTyrGlnThrArgPheLeuAr 253
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
610 AAGCTTATATTGGAACAAATATTC...TACCCACACCACTTCTTTGA 656
253 gGln...LeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleC 269
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
657 TCAATTTCTCCGCCCGAGGATGCTCCCGGAGACGGTGTGATCTCTCT 706
269 ySerAsnIleMetLeuLeuLeuGlyPheAsnThrAsnAsnMetAsn 285
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
707 GCAGCAACCCCTGTTATCATTTGTGGATTTGACATATGAACCTTGAAC 756
286 MetSerArgAlaSerValIleAlaAlaHisThrLeuAlaGlyThrSerVa 302
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
757 ATGAGTCGCTGGATGTCTATCTGTCACATAATCCAGCAGGACATCGGT 806
302 lGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgA 319
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
807 TCAGCAACGCTGCCACTGCTCCAGGCTGTTAAGTCTGGGAAGTTCCAAG 856
319 laPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
```

```
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
857 CTTTGACTGGGAAGCCAGTTCAGAACATGACACTATCATCAGAGC 906
336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
907 ATGCTTCTCTACTACACCTGACAGCATGATGGCCAATGCGCAGTGTG 956
352 pThrGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
957 GAACGGTGGCAACGACTTGTGTGGCGGACCCCTCACGATGTTGACCTTTGC 1006
369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAla 385
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1007 TTTCCAAAGCTCCCAATCTATTACCACAGAGAGATTCTCTTACAAT 1056
386 HisValAspPheIleTyrGlyLeuAspAlaProHisArgMetTyrAsnG 402
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1057 CACTTGGACTTATCTGGCCATGATGCCCTCAAGCGGTTTACAATGA 1106
402 uIleIleHisLeuMet 407
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1107 AATTGTTTCCATGATG 1122
seq_name: /cgn2.5/ptodata/1/lna/5B_COMB.seq:US-09-073-674-6
seq_documentation_block:
; Sequence 6, Application US/09073674
; Patent No. 5938189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1146 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; US-09-073-674-6
```

```
alignment_scores:
Quality: 1141.50          Length: 372
Ratio: 3.624             Gaps: 3
Percent Similarity: 84.677 Percent Identity: 56.452
```

alignment_block:

US-09-333-159-47 x US-09-073-674-6

Align seg 1/1 to: US-09-073-674-6 from: 1 to: 1146

```

37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIleL 53
|||||
19 CCACAA.....AACCTGAAGTGACCATGATATGAATGAATGATGAT 59
|||||
53 eGlnHisGlnGlyTrpCysGluGluValAlaThrGluAspGG 70
|||||
60 CACCTACTGGGATACCCAGCTGAGGAATATGAAGTTGTGACCGAAGACG 109
|||||
70 LysThrIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
|||||
110 GTTATATCTTGGGATGACAGAAATTCCTTATGGGAGGAAATTCAGAG 159
|||||
87 LysThrGlySerArgProValLeuLeuGlnHisGlyLeuValGlyL 103
|||||
160 AATATAGCGGAGACCTGTGTCATTTTGCACACGGTTTGTCTGCGCATC 209
|||||
103 ValSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleL 120
|||||
210 AGCCACAACTGGATCTCCACCTGCCCAACAACAGCTGGCCTTCATCC 259
|||||
120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
|||||
260 TGGCGGACCGCGGTACGACGTGTGGTGGGGAACAGCAGGAGGCAACACC 309
|||||
137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
|||||
310 TGGGCGAGAGAAATCTGTACTACTCGCCGACTCCGCGAAATCTGGGC 359
|||||
153 aPheSerTrpAspGluMetAlaArgPheAspLeuProAlaValIleAsnP 170
|||||
360 TTTCAGCTTTGACGAGATGGCTAAATATGACCTTCCGCCACCATGACT 409
|||||
170 HeIleLeuGlnLysThrGlyGlnGluLysIleTyTrpValGlyTrpSer 186
|||||
410 TCATCTTGAGAAACAGGAGACAGGAACTACACTAGCTTGGCCATTC 459
|||||
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
|||||
460 CAGGCGACCACTTGGTTTCTATCCCTTTTCCACCAATCCCAAGCTGC 509
|||||
203 aGlnLysIleLysMetTrpPheAlaLeuAlaProIleAlaThrValLysH 220
|||||
510 GAAACGGATCAAAACCTTCTATGATAGTCCCTGCGCACCGTGAAGT 559
|||||
220 IsAlaLysSerProGlyThrLysPheLeuLeuLeuProAspMetMetile 236
|||||
560 ACACCGAAACCTGTAAACAACTCATGCTGCTCCCTTCCTCTTC 609
|||||
237 LysGlyLeuPheGlyLysLysGluPheLeuTyTrpGlnThrArgPheLeuAr 253
|||||
610 AAGCTATATTGGAAACAAATATTC...TACCACACCACTTCTTTGA 656
|||||
253 gGln...LeuValIleTyTrpCysGlyGlnValIleLeuAspGlnIleC 269
|||||
657 TCATTTCTCGCCACCGAGGTATGCTCCCGGAGAGGTTGGATCTCTCT 706
|||||
269 ySerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnAsnMetAsn 285
|||||
707 GCAGCAACGCCCTGTTTATCATTTGTGGATTGACATATGAACATTGAAC 756
|||||
286 MetSerArgAlaSerValTrpAlaAlaHisThrLeuAlaGlyThrSerVa 302
|||||
757 ATGAGTCGCTTGGATGTGTATCTGCATATAATCCAGCAGGAACATCGGT 806
|||||
302 lGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgA 319
|||||
807 TCAGAAGCTGCTCCACTGTGTCACGCTGTTAAGTCTGGGAAGTCCCAAG 856
|||||

```

```

319 laPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
|||||
857 CTTTGTGACTGGGAAGCCAGTTCAGAACATGATGCACACTATCATCAGAGC 906
|||||
336 ThrProValArgTyArgValArgAspMetThrValProThrAlaMetTr 352
|||||
907 ATGCTCTCCTACTACACCTGACAGACATGCATGTGCCAATCGCATGTG 956
|||||
352 pThrGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
|||||
957 GAACGGTGGCAAGCACTTGTGTCGCGGACCTCAGCATGTGACCTTTGC 1006
|||||
369 euSerGluValThrAsnLeuIleTyHisLysAsnIleProGluTrpAla 385
|||||
1007 TTTCCAACTCCCAATCTCATTTACCACAGGAATTCCTCTTACAAT 1056
|||||
386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyAsnG 402
|||||
1057 CACTGTGACTTATCTGGCCATGGATGCCCTCAAGCGGTTTACAATGA 1106
|||||
402 uIleIleHisLeuMet 407
|||||
1107 AATTGTTTCCATGATG 1122
|||||
seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-227-108-1
seq_documentation_block:
; Sequence 1, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1528 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; US-08-227-108-1

```

alignment_scores:

```

Quality: 1141.50      Length: 372
Ratio: 3.624          Gaps: 3
Percent Similarity: 84.677 Percent Identity: 56.452

```

alignment_block:

US-09-333-159-47 x US-08-227-108-1

Align seg 1/1 to: US-08-227-108-1 from: 1 to: 1528

```
37 ProThrLysAlaValAspProGluAlaPheMetAsnLysSerGluLeuIle 53
19 CCCACA.....AACCCCTGAAGTGACCATGAATATAAGTCAGATGAT 59
53 eGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspG 70
60 CACCTACTGGGGATACCCAGCTGAGGATATGAAGTTGTGACCGAAAGCG 109
70 lyTyrIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
110 GTTATATCCTTGGGATCGACAGAAATTCCTTATGGGAGGAAAAATTCAGAG 159
87 LysThrGlySerArgProValValLeuLeuGlnHisGlyLeuValGlyG 103
160 AATATAGCCGCGGAGACCTGTTGCAATTTTGCACACGCTTTGCTCGCATC 209
103 yAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheLeu 120
210 AGCCACAACCTGGATCTCCACCTGCCCAACACAGCGCTGGCCCTTCATCC 259
120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
260 TGGCGGACCGCGGTACGACGTGTGGCTGGGGAACAGCAGCGGCAACACC 309
137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
310 TGGGCGAGGAGGATCTGTACTATCGCCGCGACTCCGTCGAATTCCTGGGC 359
153 aPheSerTyrAspGluMetAlaArgPheAspLeuProAlaValIleAsn 170
360 TTTCAGCTTTCACGAGATGGCTAATATGACCTCCGCGCCACCATTTGACT 409
170 heLeuLeuGlnLysThrGlyGlnGluLysIleTyrValGlyTyrSer 186
410 TCATCTTGAAGAAACCGGACGACGACGCTACACTACGTTGGCCATTC 459
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
460 CAGGCGACCATGTTGTTTCATCGCTTTCCACCAATCCCAAGCTGGC 509
203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
510 GAAAGGATCAAAACCTTCTATGATAGCTCCGCTGGCCACCGTGAAT 559
220 lAlaLysSerProGlyThrLysPheLeuLeuLeuProAspMetMetile 236
560 ACACCGAAACCCCTGTAAACAACACTCATGCTCGCTCCCTTCTCTCTTC 609
237 LysGlyLeuPheGlyLysGluPheLeuTyrGlnThrArgPheLeuAr 253
610 AGCTTATATTGGAACAAATATTC...TACCCACACCATCTCTTTGA 656
253 gGln...LeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleC 269
657 TCAATTTCTGCCCGGAGGTATGCTCCCGGAGACGCGTGTATCTCTCT 706
269 ysSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnMetAsn 285
707 GCACGACGCGCTGTTTATCATTTGTGGATTGTACACTATGAAGTGAAC 756
286 MetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSera 302
757 ATGAGTCCGCTGGATGTATCTGTCAATTAATCCAGCAGGAACATCGGT 806
302 lGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArg 319
807 TCAGACCTGCTCCACTGGTCCCGAGGCTGTTAAGTCTGGGAGGTCCAG 856
```

```
319 laPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
857 CTTTTTGACTGGGAAGCCAGTTCAGACATGATGCACATATCATCAGAGC 906
336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
907 ATGCTCTCCCTACTACAACCTGACAGACATGATGTGCCAATCGCAGTGTG 956
352 pThrGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
957 GAACGGTGGCAACGACTGTCTGGCGGACCCCTCAGATGTTGACCTTTTGC 1006
369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAla 385
1007 TTTCCAGCTCCCAATCTCATTTTACCACAGAGATTCCTCTTCAAT 1056
386 HisValaspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGl 402
1057 CACTTGGACTTATCTGGGCATGATGATGCCCTCAAGCGGTTTACAATGA 1106
402 uIleIleHisLeuMet 407
1107 AATTGTTCCATGATG 1122
seq_name: /cgn2.6/ptodata/1/lna/5b_COMB.seq:US-09-073-674-1
seq_documentation_block:
; Sequence 1, Application US/09073674
; Patent No. 598189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/09/073,674
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd W.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1528 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; US-09-073-674-1
```

alignment_scores:

Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3

61 GluGluTyrGIuIuAlaIaThrCluaSpGlyTyrlLeuSerVal...As 76
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
14769 GAGGAANTTAGATTGCAGACCGAAGATGGTTTGTTATAGATTTCGCCA 14818

76 nArgIleProArgGlyLeuValGlnProLysIsthrGlySerArgProv 93
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
14819 CTTGATACCANAATATAGAACGACAGATCTGCAGAAGAAAGAGGCCAC 14868

93 alValLeuGlnHisGlyLeuValGlyGIyAlaSerAsnTrpIleSer 109
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```
14869 CCATTTTGTACATCGCGCTTTGCAACGACAGTGGTCTGTCGCATCC 14918
110 AsnLeuProAsnAsnSerLeuGlyPheLeuLeuAlaAspAlaGlyPheAs 126
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
14919 AAT...GGTAGAAAATCTCTGGCATATTCTGTATCAATCCGTTACGA 14965
126 pValTrpMetGlyAsnSerArgGlyAsn.....AlaTrpSerArg. 139
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
14966 CATATGGTTAGGAATACAGATCGCGGTTTAGCCCGAATGGACGAAG 15015
140 ..LysHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15016 CGAAAGTACCGACACTAGCT.....TCCAGGTGGAGCTGGGAC 15053
156 TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLe 172
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15054 CTTCCGCGAGATGGTAAAGTAGATCTGACCCCTTTTGATGATACCGTGT 15103
172 uGlnLysThrGlyGlnGluLysIleTyrValGlyTyrSerGlnGlyT 189
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15104 AGCTAAGACCGAGTTGAAAGCTTACTTTGATCTCGCAATCTCAGGGCA 15153
189 hrThrMetGlyPheIleAlaPhe..... 196
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15154 CTACACAGGGGTTTATGGCTTGCTCAACGAGATAGTTTTCCTCCCTCC 15203
197 .....SerThrMetProGluLeuAlaGlnLysIleLysMetTyrPh 210
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15204 GGTTCCGGGATCTAAGAATCTTTTTCACCTTCAAGATCGCAACTATAT 15253
210 eAlaLeuAlaProIleAlaThrValLysHisAlaLysSerProGlyThrL 227
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15254 TGCCTTGGCCCC.....GCAGTGTATCTCGTCC. 15284
227 ysPheLeuLeuLeuProAspMetMetIleLys..... 237
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15285 .....TTACTTAACGAGAAATGTTGTAAAGCTTATGACAAAGAAATC 15329
238 .....GlyLeuPheGlyLysGluPheLeuTyrGlnThrArgPh 251
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15330 GAAATCCCTGGTCTTGTGGTGAACAGAGCTTTTTCAGATATATGATGAT 15379
251 eLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnI 268
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15380 TGTAAAGAACTGTGCGTT.....GTGAGAGCTTGTCTCCTTTG 15420
268 leCysSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnAsnMet 284
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15421 TTTGTACACCATCTTCAATATACCTGTTGATTGGAACGATACCTTTGG 15470
285 AsnMetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSe 301
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15471 GATACCGCATTAAGAGATCGCCATTTCTCTGTTTCGCCAGTCCCATGTTTC 15520
301 rValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuA 318
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15521 AGTGAAG...TTGATGAATGGTGGCTGTACCCGACCCCAAGGTA. 15566
318 rGAlaPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGln 334
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15567 ..AGTTTAAATTTGGTTCCTCCATAGATG.....TTC 15596
335 ProThrProValArgTyrArgValArgAspMetThrValProThrAla.. 350
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15597 CCCGACATGTAAAGTGGTTTTCAGACGCATCAAAAGGCCCAATATCTA 15646
351 .MetTrpThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysM 367
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15647 CTTGTTTGTCCAAAGCAAGATAGATTG.....GTGACCGGAGAA 15687
367 etLeuLeuSerGluValThrAsnLeuIleTyrHisLysAsn..... 380
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15688 GACTAATCAATCATTTCTGTCATGCGAGTCGATGTCACACTACAAGATC 15737
```

```
381 .....lleProGluTrpAlaHisValAspPheIleTrpGlyLeuAspAl 395
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15738 TGGTACATTGATGATGATGCCATATTGATGCTCTATGGCACATGATGT 15787
395 aProHisArgMetTyrAsnGluIleIle 404
: :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
15788 CATAGACAGAATTGTTAAACCAATTTTA. 15815
seq_name: /cgn2_6/ptodata/1/lna/5A_COMB.seq:US-08-751-782-2
seq documentation block:
; Sequence 2, Application US/08751782
; Patent No. 5821352
; GENERAL INFORMATION:
; APPLICANT: Heintz, Nathaniel
; APPLICANT: Gubbay, Johnathan
; APPLICANT: Skinner, Michael
; TITLE OF INVENTION: A cDNA Library Prepared during
; TITLE OF INVENTION: Regression of Rat Prostate and Methods of Use Thereof
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,782
; FILING DATE: 18-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson, Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; IMMEDIATE SOURCE:
; CLONE: 10.2
US-08-751-782-2
```

```
alignment_scores:
    Quality: 155.00      Length: 50
    Ratio: 3.780        Gaps: 0
Percent Similarity: 82.000 Percent Identity: 56.000
alignment_block:
US-09-333-159-47 x US-08-751-782-2 ...
Align seg 1/1 to: US-08-751-782-2 from: 1 to: 178
139 ArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSe 155
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7 AGGAAACATGTGCGCTTAGACCCAGATTCTAAAGAAATTTTGGGATTTAG 56
155 rTyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleL 172
```



```
101 lGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsn..... 114
117 GAATGGAAGAGAACTCTTCTAGAAATTTGCAATCAATTTAGAGAAG 166
115 .....SerLeuGlyPhe.....IleLeuAlaAspAla 123
167 ATTTGGAACTGGCTTCGAGATGAACTGTATCAATGGCTCAATCA 216
124 GlyPheAspValTrp.....MetGlyAsnSer.....ArgG1 134
217 GCAATCGAAATTAATCTTATGCAAGGAAATTCGGCGTCAGAACGAATGAA 266
134 yAsnAlaTrpSerArgLysHisLys.....ThrLeuS 145
267 TAATGCGAAATCCCGTCGCGAAGACCTGCTGGAAATGGAAATATCAAG 316
145 erileAspGlnAspGluPheTrpAlaPheSerTyraSpGluMetAlaArg 161
317 AAATTGAGGAAGACGATCAGATCAGTATGCTATCAGTATTCACAAA 366
162 PheAspLeuProAlaValIleAsnPheIleLeuGlnLysThrGlyGlnG1 178
367 TGTGATCTCCCTGACCTCGAAATATATACCAAAAAGATAACACGGA 416
178 uLysIleTyTrValGlyTySerGlnGlyThrMetGlyPheIleA 195
417 ATCGAAAAACGATGGAGCATCTCGGGATAATGATTCGGA..... 459
195 laPheSerThrMetProGluLeuAlaGlnLysIleLysMetTyTrPheAla 211
460 .....AAATGGAAGCTTTACTCTAGCA 480
212 LeuAlaProIleAlaThrValLysHisAlaLysSerProGlyThrLysph 228
481 GCTGGA..... 486
228 eLeuLeuLeuProAspMetMetIleLysGlyLeuPheGlyLysLysGluP 245
487 .....TTTAAATCTTTTGCACGGTGTGCGTTCGAAGCGTGATG 526
245 heLeuTyTrGlnThrArgPheLeuArgGlnLeuValIleTyTrLeuCysGly 261
527 TTCTC.....ACGAATTTGAGATGAGCTATCCGATTATACATATATG 570
262 GlnValIleLeuAspGlnIleCysSerAsnIleMetLeuLeuLeuGlyG1 278
571 AGAGTGGATGCACGGAAGATGGCTCAATGTAAGTTCTCTTGGAGC 620
278 yPheAsnThrAsn.....AsnMetAsnMetSerArgAlaSerValT 292
621 TATCAATGAGAATATGAGCTGGAATTTGTAATGTGAAGAGAGGC..... 663
292 yAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuHisTrp 308
664 .....CAATCTACGATTAGTGG 681
309 SerGlnAlaVal.....AsnSerGlyGluLeu..... 317
682 GCTGATCTATTCGAGAAAATGAATAGCCACACAGTTGATCTTATCAT 731
318 .....ArgAlaPheAspTrpGlySerGlu 325
732 TCATAATATTGAAGCTCTGATGGAGAGTGTAT 765
```

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-484-106-21

seq_documentation_block:
; Sequence 21, Application US/08484106
; Patent No. 5614618
; GENERAL INFORMATION:
; APPLICANT: STILLMAN, Bruce
; APPLICANT: BELL, Stephen P

```
APPLICANT: KOBAYASHI, Ryuji
APPLICANT: RINE, Jasper
APPLICANT: FOSS, Margit
APPLICANT: McNALLY, Francis J
APPLICANT: LAURENSEN, Patricia
APPLICANT: HERSKOWITZ, Ira
APPLICANT: LI, Joachim J
APPLICANT: GAVIN, Kimberly
TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSER: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,106
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oshman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 1676 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 13..1302
US-08-484-106-21
```

alignment_scores:
Quality: 105.00 Length: 278
Ratio: 0.802 Gaps: 13
Percent Similarity: 47.122 Percent Identity: 21.942

alignment_block:
US-09-333-159-47 x US-08-484-106-21 ..
Align seg 1/1 to: US-08-484-106-21 from: 1 to: 1676

```
85 ProLysLysThrGlySerArgProValValLeuLeuGlnHisGlyLeuVa 101
111.....CAGAAAAAGAC 116
85 CCAGAAAAAGAGGATCCAGA.....CAGAAAAAGAC 116
101 lGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsn..... 114
117 GAATGGAAGAGAACTCTTCTAGAAATTTGCAATCAATTTAGAGAAG 166
115 .....SerLeuGlyPhe.....IleLeuAlaAspAla 123
167 ATTTGGAACTGGCTTCGAGATGAACTGTATCAATGGCTCAATCA 216
124 GlyPheAspValTrp.....MetGlyAsnSer.....ArgG1 134
217 GCAATCGAAATTAATCTTATGCAAGGAAATTCGGCGTCAGAACGAATGAA 266
```

```
134 yAsnAlaTrpSerArgLysHisLys.....ThrLeus 145
:||||| :||||| :|||
267 TAATCGAAATCCCGTCTCGAAGACGTGCTGGAATGAATACTAGAG 316
:||||| :||||| :|||
145 erIleAspGlnAspGluPheTrpAlaPheSerTyrAspGluMetAlaArg 161
:||||| :||||| :|||
317 AAATTGAGGAACAGCATGAGATCATGATGATGATGATGATGATGATGAT 366
:||||| :||||| :|||
162 PheAspLeuProAlaValIleAsnPheIleLeuGlnLysThrGlyGlnG 178
:||||| :||||| :|||
367 TGTGATCTCCCTGGACTTCGAAATATATATACCAAAAAGATACACCGGA 416
:||||| :||||| :|||
178 uLysIleTyrTyrValGlyTyrSerGlnGlyThrThrMetGlyPheIleA 195
:||||| :||||| :|||
417 ATTCGAAAACGATTGGAGCATCTCGCGGATAATGATTTCGGA..... 459
:||||| :||||| :|||
195 laPheSerThrMetProGluLeuAlaGlnLysIleLysMetTyrPheAla 211
:||||| :||||| :|||
460 .....TTATATATCTTTTCACGGTGTGCGTTCGAAGCGTGATG 480
:||||| :||||| :|||
212 LeuAlaProIleAlaThrValIleHisAlaLysSerProGlyThrLysPh 228
:||||| :||||| :|||
481 GCTGGA..... 486
:||||| :||||| :|||
228 eleuLeuLeuProAspMetMetIleLysGlyLeuPheGlyLysLysGluP 245
:||||| :||||| :|||
487 .....TTATATATCTTTTCACGGTGTGCGTTCGAAGCGTGATG 526
:||||| :||||| :|||
245 heLeuTyrGlnThrArgPheLeuArgGlnLeuValIleTyrLeuCysGly 261
:||||| :||||| :|||
527 TTCTC.....ACAGATTTCAGAAATGAGCTATCGGATTATACATATATG 570
:||||| :||||| :|||
262 GlnValIleLeuAspGlnIleCysSerAsnIleMetLeuLeuGlyG 278
:||||| :||||| :|||
571 ACAGTGGATGACCGGAAAGATGGCGTCAATGAAAGTCTCTTGGAGC 620
:||||| :||||| :|||
278 yPheAsnThrAsn.....AsnMetAsnMetSerArgAlaSerValT 292
:||||| :||||| :|||
621 TATCAATGAGATATGAAGTCAATGTAATGTGAAGAGAGCC..... 663
:||||| :||||| :|||
292 yAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuHisTrp 308
:||||| :||||| :|||
664 .....CAATCTACGATTAGTTGG 681
:||||| :||||| :|||
309 SerGlnAlaVal.....AsnSerGlyGluLeu..... 317
:||||| :||||| :|||
682 GCTGATCTATTCCGAGAAAATGAATAGACCAACAGTTGATTCTTATCAT 731
:||||| :||||| :|||
318 .....ArgAlaPheAspTyrGlySerGlu 325
:||||| :||||| :|||
732 TGATAATATTGAGCTCTGATTGGAGAGTGAT 765
:||||| :||||| :|||
seq_name: /cgn2_6/ptodata/1/ins/5B_COMB.seq:US-08-602-359A-24
seq_documentation_block:
; Sequence 24, Application US/08602359A
; Patent No. 5942430
; GENERAL INFORMATION:
; APPLICANT: ROBERTSON, Daniel E.
; APPLICANT: MURPHY, Dennis
; APPLICANT: REID, John
; APPLICANT: MAFFIA, Anthony
; APPLICANT: LINK, Steven
; APPLICANT: SWANSON, Ronald V.
; APPLICANT: WARREN, Patrick V.
; APPLICANT: KOSMOTKA, Anna
; TITLE OF INVENTION: ESTERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 EXECUTIVE SQUARE, STE 1400
; CITY: LA JOLLA
; STATE: CALIFORNIA
```

```
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,359A
FILING DATE: February 16, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HAILE, LISA A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09010/010001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1041 NUCLEOTIDES
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: GENOMIC DNA
US-08-602-359A-24

alignment_scores:
Quality: 102.00 Length: 180
Ratio: 1.052 Gaps: 8
Percent Similarity: 53.889 Percent Identity: 22.778

alignment_block:
US-09-333-159-47 x US-08-602-359A-24
Align seg 1/1 to: US-08-602-359A-24 from: 1 to: 1041

20 LeuIleLeuValAlaTyrMetPheGlnArgAsnValAsnSerValHisMe 36
||||| :||||| :|||
178 CTGCTCTGGCGCTGTGCTGCTCTGGCGCTGTGCTGCTGCTGCTGCTGCT 227
36 tProThr.....LysAlaValAspProGluAlaPheMetAsnI 49
:||||| :||||| :|||
228 CCGCGCTTATGTGGCTATAGTAGTCTGCACCGGAGAGCTGCGCCCG 277
49 leSerGluIleIleGlnHisGlnGlyTyrProCysGluGluTyrGluVal 65
:||||| :||||| :|||
278 TTGCGCCG.....GAGGGCTCACCTACAAAGAGTTCACGCTG 315
66 AlaThrGluAspGlyTyrIleLeuSerValAsnArgIleProArgGlyLe 82
:||||| :||||| :|||
316 ACCGCGGAGGATGCTTGTGCTG.....CGGGCTG 347
82 uValGlnProLysLysThrGlySerArgProValValLeuLeuGlnHisG 99
||| :||||| :|||
348 GGTGCTGGGCCCCCGCGCTGGGGGCAACCGGTGTTCTGTTTGTATGACG 397
99 lyLeuValGly...GlyAlaSerAsnTrpIleSerAsnLeuProAsn 114
|| :||||| :|||
398 GGTATCTGCTGGCTCGCGCTCGCGCCCTACATGCTGTGCTGGCCCGGAG 447
115 SerLeuGlyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGlyAs 131
||||| :||||| :|||
448 .....CTCGTGGAGTGGGGGTACCCCGGTGTTGTTGTTGTTGTTG 482
131 nSerArgGlyAsnAlaTrpSerArgLysHisLysThrLeuSerIleAspG 148
:||||| :||||| :|||
483 CTTCCGGGCCACGCGGAGAGCGGGGCTCGACGAGGATTGGG..... 525
```

```

148 InAspGluPheTrpAlaPheSerTyrAspGluMetAlaArgPheAspLeu 164
526 .....CCCCGGGAGGTGCTGGATGCC 546
165 ProAlaValIleAsnPheIleLeuGlnLysThrGlyGlnGluLysIleTyr 181
547 CGGGCTGTGGGGGTATGCTCGGAGGGGTCCCGGCCCGGGGATAT 596
181 rTyrValGlyTyrSerGlnGlyThrThrMetGlyPheIle 194
597 ATTGGTGGGTTCAGTATGGCGGCGCTGTAGCGATCGT 636

```

seq_name: /cqn2_6/ptodata/1/lna/5A_COMB.seq.us-08-232-519-1

```

seq_documentation_block:
; Sequence 1, Application US/08232519
; Patent No. 5484725
; GENERAL INFORMATION:
; APPLICANT: Kageyama, Bunji
; APPLICANT: Nakae, Masanori
; APPLICANT: Yagi, Shigeo
; TITLE OF INVENTION: NORBORANE TYPE ESTER HYDROLASE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,519
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-96286
; FILING DATE: 22-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr, James F
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: SHGN-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1167 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Acetobacter pasteurianus
; STRAIN: ATCC 12873
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1164
; US-08-232-519-1

```

```

alignment_scores:
  Quality: 98.50      Length: 400
  Ratio: 0.518      Gaps: 20
  Percent Similarity: 47.500  Percent Identity: 20.250

```

alignment_block:
US-09-333-159-47 x US-08-232-519-1

```

Align seg 1/1 to: US-08-232-519-1 from: 1 to: 1167
38 ThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIleIleG1 54
250 ACGGCAGATGTGGATATTGAAGCCCTTATC..... 279
54 nHisGlnGlyTyrProCysGluGluTyrGlu.....ValAlaThrGluA 69
280 ....AAAACCTTCATGCAGAAAAACCCACAGGATCGCGCTCAACAGAAG 325
69 spGlyTyrIleLeuSerValAsnArgIleProArgGly.....LeuVal 83
326 ATGCTCTGCGCGGGAACCCCAAGCAGGTACGGTAGGCGGAACACACGTA 375
84 GlnProLysLysThrGlySerArg.....ProValValLeuLeuG1 97
376 AATGTGCGTGATGTGGCAGCAGGAGGCGCCCATTTGCTGGTG... 423
97 nHisGlyLeuValGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnA 114
424 .CACGGITTTGGCGGAGATATCAGCACTGGCTG.....CTCACACAGG 466
114 snSerLeuGlyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGly 130
467 ATGCTTGGCGGCAGAAAAAGCGCGTAATTGGCTTTGATCTG..... 507
131 AsnSerArgGlyAsnAlaTrpSerArgLysHis.....LysThrLe 144
508 .....SCGGGGCATGGGGCTTCCTCTAAACACGTGGGCACAGCACGCT 551
144 uSerIleAspGlnAspGluPheTrpAlaPheSerTyrAspGluMetAla 161
552 GGCCTT..... 558
161 rgPheAspLeuProAlaValIleAsnPheIleLeuGlnLysThrGlyGln 177
559 .....TTGGCGGGTGGTGAAGCGAATTGCTGCAAAACCTTAAAAATA 600
178 GluLysIleTyrValGlyTyrSerGlnGlyThrThrMetGlyPheI1 194
601 GAAAAAGCCCATGTGGTGGCCATTTCTTTGGGG.....GGCGGCAT 641
194 eAlaPheSerThrMetProGluLeuAlaGlnLysIleLysMetTyrPheA 211
642 TGCCCTGACCTGCTGCGAGATCACCTGATCAGGTTCCAGCCTGAACC 691
211 laLeuAlaProIleAlaThrValLysHisAlaLysSerProGlyThrLys 227
692 TTTTGGCCCCAGCGCGGTG.....GGAAGGATGTGAATGCAGAT 732
228 PheLeuLeuLeuProAspMetMetIleLysGlyLeuPheGlyLysLysG1 244
733 TTTATC.....AGCCCATTTGTGGATAGTGA 758
244 uPheLeuTyrGlnThrArgPheLeuArgGlnLeuValIleTyr..... 258
759 AAGCAGCGCGCATATGAAGGCTGTTTTCAAATGCTGCTATATAACAAG 808
259 ..LeuCysGlyGlnValIleLeuAspGlnIleCysSerAsnIleMetLeu 274
809 CCCTAGTGGCGCGTAAGATGGTGGATGCCGTG..... 840
275 LeuLeuGlyGlyPheAsnThrAsnAsnMetAsnMetSerArgAlaSerVa 291
841 .....CTGCGTGCAGCTAGGCTAGATGGCGCGCGGGAT..... 873
291 lTyrAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuHist 308
874 ....GCCTGCACGCTATTGCTAAAGCGTGTCTCCCAAC..... 909
308 rpSerGlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpGlySer 324

```

910GGGCATCAGCGGATGATCTGCACCTCGGTGCTAGCTGGGGCG 951
325 GluThrLysAsnLeuGluLysCysAsnGlnProThrProValArgTyrAr 341
|||||
952 GAACA..... 957
341 gValArgAspMetThrValProThrAlaMetTrpThrGlyGlnAspT 358
||||| : : : : :
958CCTACCCAGATTCTGGGGCAAGGAAGATG 988
358 rpLeuSerAsnProGluAspValLysMetLeuLeuSerGluValThrAsn 374
: : : : :
989 AAATCTTCTGTCTCCAAACCGCGCTGCCCTGCCAGATGTCATCCCCGTG 1038
375 LeuIleTyrHisLys.....AsnIleProGluTrpAlaHisValAspPh 389
: : : : :
1039 ACAGTGTATGAAGAAACACAGGCCATCTGCCGAGCTTGAACATGCAACAGA 1088
389 eIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGluIleHisL 406
: : : : :
1089 TGTG.....AACAAAGCCATTGCC 1108
406 eu...MetGlnGlnGluThrAsnLeuSerGlnGlyArgCysGluAla 421
: : : : :
1109 TGTGTAAGACCCCGAAGCGCGCTGAGCATGGCCCGGATGACGCG 1158

OM of: US-09-333-159-47 to: EST:* out_format : pfs

Date: Sep 1, 2001 10:13 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet_p2n.model -DEV=rlp
-Q-CG2_1/USP10.spool/US09333159/runat_310822001_122330_1373/app_query.fasta_1.488
-DB=EST -FMT=fastap -SUFFIX=p2n.rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOPCL=0.000 -LOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000
-GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blos62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=200000000 -USER=US09333159 -CGN1_1_3139 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-333-159-47

Query length: 423

Database: EST:*

Database sequences: 10238115

Database length: 431459454

Search time (sec): 1357.660000

score_list:	Strd Orig	ZScore	EScore	Len	Documentation
gb_htc:AK019504	+ 1836.00	3851.43	2.1e-205	2927	AK019504 Mus musculus 0 day ne
gb_htc:AK009571	+ 1139.50	2383.04	1.3e-123	1349	AK009571 Mus musculus adult ma
gb_htc:AK009875	+ 1139.50	2383.04	1.3e-123	1349	AK009875 Mus musculus adult ma
gb_htc:AK010093	+ 1139.50	2383.03	1.3e-123	1350	AK010093 Mus musculus adult ma
gb_htc:AK010139	+ 1139.50	2383.03	1.3e-123	1350	AK010139 Mus musculus adult ma
gb_htc:AK009431	+ 1139.50	2383.01	1.3e-123	1352	AK009431 Mus musculus adult ma
gb_htc:AK009479	+ 1138.50	2380.97	1.7e-123	1343	AK009479 Mus musculus adult ma
gb_htc:AK009523	+ 1138.50	2380.95	1.7e-123	1345	AK009523 Mus musculus adult ma
gb_htc:AK009729	+ 1138.50	2380.94	1.7e-123	1346	AK009729 Mus musculus adult ma
gb_htc:AK010116	+ 1138.50	2380.93	1.7e-123	1347	AK010116 Mus musculus adult ma
gb_htc:AK010058	+ 1138.50	2380.92	1.7e-123	1348	AK010058 Mus musculus adult ma
gb_htc:AK009300	+ 1138.50	2380.91	1.7e-123	1349	AK009300 Mus musculus adult ma
gb_htc:AK009474	+ 1138.50	2380.91	1.7e-123	1349	AK009474 Mus musculus adult ma
gb_htc:AK009546	+ 1138.50	2380.91	1.7e-123	1349	AK009546 Mus musculus adult ma
gb_htc:AK009573	+ 1138.50	2380.91	1.7e-123	1349	AK009573 Mus musculus adult ma
gb_htc:AK009459	+ 1138.50	2380.91	1.7e-123	1350	AK009459 Mus musculus adult ma
gb_htc:AK009525	+ 1138.50	2380.91	1.7e-123	1350	AK009525 Mus musculus adult ma
gb_htc:AK009773	+ 1138.50	2380.91	1.7e-123	1350	AK009773 Mus musculus adult ma
gb_htc:AK010061	+ 1138.50	2380.91	1.7e-123	1350	AK010061 Mus musculus adult ma
gb_htc:AK010125	+ 1138.50	2380.91	1.7e-123	1351	AK010125 Mus musculus adult ma
gb_htc:AK009428	+ 1138.50	2380.90	1.7e-123	1351	AK009428 Mus musculus adult ma
gb_htc:AK010019	+ 1138.50	2380.85	1.7e-123	1356	AK010019 Mus musculus adult ma
gb_htc:AK010035	+ 1138.50	2380.85	1.7e-123	1356	AK010035 Mus musculus adult ma
gb_htc:AK010203	+ 1136.50	2376.66	2.9e-123	1350	AK010203 Mus musculus adult ma
gb_htc:AK010124	+ 1135.50	2374.63	3.8e-123	1340	AK010124 Mus musculus adult ma
gb_htc:AK009413	+ 1135.50	2374.59	3.8e-123	1345	AK009413 Mus musculus adult ma
gb_htc:AK010236	+ 1135.50	2374.57	3.8e-123	1347	AK010236 Mus musculus adult ma
gb_htc:AK009930	+ 1134.50	2372.43	5.0e-123	1349	AK009930 Mus musculus adult ma
gb_htc:AK009537	+ 1133.50	2370.30	6.6e-123	1350	AK009537 Mus musculus adult ma
gb_htc:AK009359	+ 1131.50	2368.21	8.6e-123	1346	AK009359 Mus musculus adult ma
gb_htc:AK009544	+ 1131.50	2366.10	1.1e-122	1345	AK009544 Mus musculus adult ma
gb_htc:AK010103	+ 1130.50	2363.96	1.5e-122	1347	AK010103 Mus musculus adult ma
gb_htc:AK009560	+ 1129.50	2361.85	1.9e-122	1346	AK009560 Mus musculus adult ma
gb_htc:AK010026	+ 1129.50	2361.82	2.0e-122	1349	AK010026 Mus musculus adult ma
gb_htc:AK010231	+ 1128.50	2359.71	2.6e-122	1348	AK010231 Mus musculus adult ma
gb_htc:AK009437	+ 1127.50	2357.60	3.4e-122	1346	AK009437 Mus musculus adult ma
gb_htc:AK010036	+ 1125.50	2353.34	5.8e-122	1348	AK010036 Mus musculus adult ma
gb_htc:AK010148	+ 1124.50	2351.21	7.6e-122	1349	AK010148 Mus musculus adult ma
gb_htc:AK010106	+ 1121.50	2344.85	1.7e-121	1349	AK010106 Mus musculus adult ma
gb_htc:AK010173	+ 1120.50	2342.75	2.3e-121	1346	AK010173 Mus musculus adult ma

gb_htc:AK010123 + 1110.50 2321.50 3.4e-120 1350 AK010123 Mus musculus adult
gb_htc:AK009923 + 1109.50 2318.87 4.8e-120 1407 AK009923 Mus musculus adult
gb_htc:AK010091 + 1086.50 2270.60 2.4e-117 1347 AK010091 Mus musculus adult
gb_htc:AK010110 + 1075.50 2247.21 4.7e-116 1353 AK010110 Mus musculus adult
seq_name: gb_htc:AK019504

seq_documentation_block: 2927 bp mRNA HTC 08-FEB-2001
LOCUS AK019504
DEFINITION Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched
library, clone:4632427C23, full insert sequence.

ACCESSION AK019504
VERSION AK019504.1 GI:12859754

KEYWORDS CAP trapper.

SOURCE Mus musculus (strain:C57BL/6J) 0 day neonate skin cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library
clone:4632427C23.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (sites)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Methods Enzymol. 303, 19-44 (1999)

2 (sites)

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itom, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

3 (sites)

Shibata, K., Itoh, M., Aizawa, K., Nageoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M.,

Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T.,

Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T.,

Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iwata, M., Ohara, E.,

Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,

Matsuura, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

4 (sites)

The RIKEN Genome Exploration Research Group Phase II Team and

FAITOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 585-590 (2001)

5 (bases 1 to 2927)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,

Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M.,

Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F.,

Imotani, K., Ishii, Y., Itoh, M., Iwata, M., Kato, H., Kawai, J.,

Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T.,

Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y.,

Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H.,

Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,

Souabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,

Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A.,

Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,

fax: 81-45-503-9216)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for

further details.

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia, Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

```

alignment_scores:      Quality: 1836.00      Length: 682
                       Ratio: 4.567      Gaps: 2
                       Percent Similarity: 58.944      Percent Identity: 54.106

alignment_block:
US-09-333-159-47 x AK019504      ..

Align seg 1/1      to: AK019504      from: 1      to: 2927

1 MetLeuGluThrLeuSerArgGlnTrpIleValSerHisArgMetGluMe 17
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
153 ATGTCAGAAATCTGTCAGAGCTGTGGACACTGTTTCGCACAGAGTGGAGAT 202

17 tTrpLeuLeuIleLeuValAlaTyrMetPheGlnArgAsnValAsnSerV 34
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
203 ATGGCTTCATGATTCTGATACGCTATTATTCATCCAAAGAAATCTGAACTCGG 252

34 aLHISMetProThrLysAlaValAspProGluAlaPheMetAsnIleSer 50
  |||::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
253 GACATTGGCCACGAAGCTCGCGATCCAGAGCATTCATGAATGTTAGC 302

51 GluIleIleGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaIaTh 67
  |||::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
303 GAAATCATCAACACACAAGGTTATCCCACTGAGGAGTATGAAGTTGCAAC 352

67 rGluaspGlyTyrIleLeuSerValAsnArgIleProArgGlyLeuValG 84
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
353 GGAAGATGGGTATCATCTCTTCTGTGAACAGAAATCCCTCGGGGGACACAC 402

84 InProLysLysThrGlySerArgProValValLeuLeuGlnHisGlyLeu 100
  :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

155 .....SerTyrAspG1 158
1353 GTTTTATCATTCTCGAATAACTTTCTCTTTCCCTTTGTATGATGATGA 1402
158 uMetAlaArgPheAspLeuProAlaValIleAsnPheIleLeuGlnLysT 175
1403 AATGCTAGTTGACCTCCAGCTGCTGATAAATTTATCTTACAGAAAA 1452
175 hrGlyGlnGluLysIleTyrValGlyTyrSerGlnGlyThrMet 191
1453 CGGGCCAGAAAAAGCTATTATGTCGGCTACTCACGGCCACCACCATG 1502
192 GlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLysIleLysMe 208
1503 GGCCTTATTGCAATTTCCACAAATGCCAGAGTAGCTCATAAATCAAAAT 1552
208 tTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLysSerProG 225
1553 GTATTTGCTTAGCTCTCTATAGCCACTGTTAAATATGCAAGAACTCTG 1602
225 lyThrLysPheLeuLeuLeuProAspMetMetIleLysGlyLeuPheGly 241
1603 GTACCAAAATTTCTGCTGCTGCCAGATATGATGATCAAGTATTATTGCG 1652
242 LysLysGluPheLeuTyrGlnThrArgPheLeuArgGlnLeuValIleTy 258
1653 AGACAAGAGTTTTTATACCACTAGATTTTTCAGACAGCTTTTATTATTA 1702
258 rLeuCysGlyGlnValIleLeuAspGlnIleCysSerAsnIleMetLeuL 275
1703 CTTTTGTGCGCAGATGATCTTGACCAATCTGCAGCAACATCATCTTAC 1752
275 euLeuGlyGlyPheAsnThrAsnAsnMetAsnMetSerArgAlaSerVal 291
1753 TCCTGGGAGGTTTTAACACAAACATATGAACATGAGCAGACAAATGTG 1802
292 TyrAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuHisTr 308
1803 TATGTCGCCCATACACCTGCTGGGACGCTGTGCAGAACATCTCCACTG 1852
308 pSerGlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpGlySerG 325
1853 GAGCCAGGCGGTGAATCTGGGGAACCTTCGTGCTTTGACTGGGAAGTG 1902
325 luThrLysAsnLeuGluLysCysAsnGlnProThrProValArgTyrArg 341
1903 AGACCAAAAAATCAGGAGAAATGTAATCAGCCAACTCTTATAGGTACAAA 1952
342 ValArgAspMetThrValProThrAlaMetTyrThrGlyGlyGlnAspTr 358
1953 GTTCGAGATATGATGCTCCCAACAGCAATGTGGACTGGAGGTCAAGACTG 2002
358 pLeuSerAsnProGluAspValLysMetLeuLeuSerGluValThrAsnL 375
2003 GCTTTCAATCCAGATGATGTGAAACATTTACTTTCTGAAGTAAACCAAC 2052
375 euIleTyrHisLysAsnIleProGluTyrPalaHisValAspPheIleTrp 391
2053 TCATCTACCACCAAGAACATTTCTGAGTGGCTCATGTGGATTTTCATCTGG 2102
392 GlyLeuAspAlaProHisArgMetTyrAsnGluIleIleHisLeuMetG1 408
2103 GGGCTGGATGCCCCCTCAGGCTGTTTACAAATGAAATATATACATCTGATGA 2152
408 nGlnGluThrAsnLeuSerGlnGlyArgCysGluAlaValLeu 423
2153 GCAG...GAGCCCAACCTTCCCCAGGACCTGCGAGGGTCAATTG 2195
seq_name: gb_htc:AK009571
seq_documentation_block:
LOCUS AK009571 1349 bp mRNA HTC 08-FEB-2001

```

DEFINITION

ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

Host: SOLR.

FEATURES

source

Location/Qualifiers

```
1..1349
  /organism="Mus musculus"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /db_xref="MGD:MGI:1895763"
  /db_xref="MGD:MGI:1914967"
  /clone="2310032G08"
  /sex="male"
  /tissue="tongue"
  /clone_lib="RIKEN full-length enriched mouse cDNA library"
  /dev_stage="adult"
52..1239
  /note="putative"
  /codon_start=1
  /protein_id="BAB26368.1"
  /db_xref="GI:12844448"
  /translation="MWLLVTVSLVSAFGGAHGLFGKLPKNPEANNVSMQIMTYWGYP
  SEEVYEDGYTLGYRIPYGGKNSENIGKRPVAYLQHLIASATNWTNLPNNSLA
  FILADAGYDVLGNSRGTWSRKNVYSPDSVEFWAFSDEMAKYDLPATIDFIVQKT
  GQSKIHVGHSGQTTIGTAFSTNPALAKIKRFYALAPVATVKYTESPKKISLIPK
  FLKLVIFGNKMPHNYLDQFLGTEVCSRELLDLCSNALFIECGEDKKNLNVRFDV
  YLGNFAGSTQDLFWAQLAKSGKIQATNWSPLQNMLHYNQKTPPYTDVSAWTVPI
  AVNGHDIILADPDQVAMLLPLNLLYHKEILPYNHLDFIWAMDAPQEVYNEIVTM
  AED"
```

CDS

```
BASE COUNT      396 a  290 c  284 g  379 t
ORIGIN

alignment_scores:
  Quality: 1139.50      Length: 411
  Ratio: 3.485          Gaps: 4
  Percent Similarity: 79.562  Percent Identity: 52.555

alignment_block:
US-09-333-159-47 x AK009571
Align seg 1/1 to: AK009571 from: 1 to: 1349
```

```
7  ArgGlnTrrPileValSerHisArgMetGluMetTrpLeuLeuLeuVa 23
|||||
34  AGGCACCTGG.....CACTCAAGATGTGGCTGCTATTAGTAAC 71
|||||
23  lAlaTyrMetPheGln...ArgAsnValAsnSerValHisMetProThrL 39
|||||
72  AAGTGTCTATCTGCTATTTGGAGTGCACATGGCCTATTGGAAACTGG 121
|||||
39  yAlaValAspProGluAlaPheMetAsnIleSerGluIleLeGlnHis 55
|||||
122  GTCCCAAAACCCCTGAAGCAACATGAATGTAGTCAGATGATGAACCTTAC 171
|||||
56  GlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspGlyTyrIl 72
|||||
172  TGGGATATCCAAAGTGAGGAATATGAAGTTGTACTGAAGATGGCTACAT 221
|||||
72  eLeuSerValAsnArqileProArgGlyLeuValGlnProLysLysThrG 89
|||||
222  TCTGGGGTCTATAGAAATTCCTATTTGGGAGAGAAATTCAGAGATATCG 271
|||||
89  lYSerArgProValValLeuLeuGlnHisGlyLeuValGlyGlyAlaSer 105
|||||
272  GCAAGAGACCTGTGGCATATTTCAGCATGGTTTGTATTGCATCAGCCACA 321
|||||
106  AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs 122
|||||
322  AACTGGATTACAAATCTGCCAAACACAGCCCTGGCCCTTCATTCTAGCAGA 371
|||||
122  pAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerA 139
|||||
372  TGCTGGCTATGATGTGGCTGGGACACAGTCGAGGGAATACATGGTCCC 421
```

```
139  rgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
|||||
422  GGAATAATGTATACTATTATCCACGACAGTCAGTTGAATTCTGGCTTCAGC 471
|||||
156  TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLe 172
|||||
472  TTTGATTAATGGCTAATAATAGACTCTCCACGCCACCATAGACTTCATTGT 521
|||||
172  uGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSerGlnGlyT 189
|||||
522  ACAGAAACATGACAGAGAGATACACTATGTTGGTCACTCTCAGGSCA 571
|||||
189  hrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
|||||
572  CCATATCGGTTTATTGGCTTTTCTACCAATCCTGCTCTGGCTAAAGAAA 621
|||||
206  IleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLy 222
|||||
622  ATCAAGAGGTTTATGATTAAGTCCAGCTGCTACTGTGAAGTATACAGA 671
|||||
222  sSerProGlyThrLysPheLeuLeuLeuProAspMetMetIleLysGlyL 239
|||||
672  AAGTCCCTTTAAAAAGATTTTCACATTATTCCTAAGTTTCTTCTCAAGGTGA 721
|||||
239  euPheGlyLysLysGluPheLeu.....TyrGlnThrArgPheLeu 252
|||||
722  TATTGTGTAACAAAATGTTCTACGCCCACTACTTAGATCAATTTCTT 771
|||||
253  ArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleCy 269
|||||
772  GGTACGGGAAGTG.....TGCTACGGGAGCTGTAGATCTTCTCTG 812
|||||
269  sSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetAsnM 286
|||||
813  CAGCAACGCTTATTCATCTTCTGTGGATTGACAAAGAAACAACTTAATG 862
|||||
286  etSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVal 302
|||||
863  TGAGTCGCTTTGATGTATCTAGGCATATATCCAGCAGGAACATCTACT 912
|||||
303  GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgAl 319
|||||
913  CAAGACCTTTCCACTGGGCACAGCTTCTAAATCTGGGAGCTTCAAGC 962
|||||
319  aPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnProT 336
|||||
963  CTATAACTGGGGAAGTCCATACAGAACATGTTACACTACAAATCAGAAA 1012
|||||
336  hrProValArgTyrArgValArgAspMetThrValProThrAlaMetTrp 352
|||||
1013  CGCCTCCCTACTATGATGTGTGAGCCATGACCGTGCCCAATTGACGTGG 1062
|||||
353  ThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuLe 369
|||||
1063  AACGGIGCCCATGACATCTCGCTGATCCCAAGATGTGGCAATGCTGCT 1112
|||||
369  uSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAlaH 386
|||||
1113  TCCCAAACTCCCAACCTTCTGTACATAAGGAGATTTCTCCCTACAATC 1162
|||||
386  lValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGlu 402
|||||
1163  ACCTGGACTTCTATCTGGGCGATGGTGGCCCTCAAGAGGTTTACAATGAG 1212
|||||
403  IleIleHisLeuMetGlnGlnGluThrAsn 413
|||||
1213  ATAGTTACCATGATGGCAGACACTAACAGAAAT 1245
```

seq_name: gb_htc:AK009875

seq_documentation_block:

LOCUS AK009875 1349 bp mRNA

DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched

HTC

08-FEB-2001

AK010093	
AK010093.1	GI:12845297
CAF trapper.	
Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,	
clone-lib:RIKEN full-length enriched mouse cDNA library	
clone:2310067K20.	
Mus musculus	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;	
1 (sites)	
Carninci,P. and Hayashizaki,Y.	
High-efficiency full-length cDNA cloning	
Methods Enzymol. 303, 19-44 (1999)	
2 (sites)	
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,	
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.	
Normalization and subtraction of cap-trapper-selected cDNAs to	
prepare full-length cDNA libraries for rapid discovery of new genes	
Genome Res. 10 (10), 1617-1630 (2000)	
20499374	
3 (sites)	
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,	
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,	
Hakuchi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T.,	
Kurada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T.,	
Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,K., Izawa,M., Ohara,E.,	
Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,	
Matsura,S., Okazaki,Y., Muramatsu,M., Inoue,Y. and Hayashizaki,Y.	
RIKEN integrated sequence analysis (RISA) system - 384-format	
sequencing pipeline with 384 multicapillary sequencer	
Genome Res. 10 (11), 1757-1771 (2000)	
20530913	
4 (sites)	
The RIKEN Genome Exploration Research Group Phase II Team and	
FANTOM Consortium.	
Functional annotation of a full-length mouse cDNA collection	
Nature 409, 685-690 (2001)	
5 (bases 1 to 1350)	
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,	
Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,	
Haragaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiroka,T., Hori,F.,	
Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawaji,J.,	
Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,	
Miyazaki,A., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Okazaki,Y.,	
Okido,T., Owa,C., Saio,H., Saio,R., Sakai,C., Sakai,K., Sano,H.,	
Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,	
Socabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,	
Tanaka,T., Tejjima,Y., Toyota,T., Yamamura,T., Yasunishi,A.,	
Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.	
Direct Submission	
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of	
Physical and Chemical Research (RIKEN), Laboratory for Genome	
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),	
Riken Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,	
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,	
URL: http://genome.gsc.riken.go.jp/ , Tel:81-45-503-9222,	
Fax:81-45-503-9216)	
Please visit our web site (http://genome.gsc.riken.go.jp/) for	
further details.	
cDNA library was prepared and sequenced in Mouse Genome	
Encyclopedia Project of Genome Exploration Research Group in Riken	
Genomic Sciences Center and Genome Science Laboratory in RIKEN.	
Division of Experimental Animal Research in Riken contributed to	
prepare mouse tissues. First strand cDNA was primed with a primer	
[5'-GAGAGAGAGGATCCAGTCTTTTTCCTTTTTTTTNN 3'], cDNA was	
prepared by using trihaloetho thermo-activated reverse transcriptase	
and subsequently enriched for full-length by cap-trapper. cDNA went	
through one round of normalization to Rot = 5.0 and subtraction to	
Rot = 25.0. Second strand cDNA was prepared with the primer adapter	
of sequence [5'	
GAGAGAGATTCCTCGATTAAATTAATTCCTCCCCCCCCCCC 3']. cDNA was cleaved	
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.	
Host: SOLR.	

seq_documentation_block:			
LOCUS	AK010093	1350 bp	mRNA
DEFINITION	Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310067k20, full insert sequence.	HTC	08-FEB-2001


```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGD:MGI:1894047"
/db_xref="MGD:MGI:1914967"
/clone="2310021816"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
55.1242
/notes="putative"
/codon_start=1
/protein_id="BAB26283.1"
/db_xref="GI:12844223"
/translation="MWLLVTSVLSAFGAHGLFGKLPKNPEANNVVSQMITYWGYP
SEYEVVTDGVTILGYRIPYIGKKNSEINIGKRPVAYLOHGLIASATNWTINLPNNSLA
FILADAGVWLGNRSNTWSRKNVYSPDSVEFAFSDEMAYDLPATIDFIVOKT
GOEKIHVGHSGOCTTIGFTAFSTNPALAKIKRFYALAPVATVKYTESPKKISLPK
FLKVIYFGKKNFPHNYLDOFLGTEVCSRELLDLCSNALFICGDFDKKLNUNSRIDV
YLGNFAGTSTQDLFHWALQSKGKLQAYNMGSPQLNMLHYNOKTPPYDYDVSAINTVPI
AVWNGHDLADPDQVAMLLPKLPNLLYHKEILPYNHLDFIWAMDAPQEVYNEIVTMM
AED"
BASE COUNT      397 a 291 c 285 g 379 t
ORIGIN

```

```

alignment_scores:
  Quality: 1139.50      Length: 411
  Ratio: 3.495          Gaps: 4
  Percent Similarity: 79.319 Percent Identity: 52.555

alignment_block:
US-09-333-159-47 x AK009431 ..
Align seg 1/1 to: AK009431 from: 1 to: 1352

7  ArgGlnTrpIleValSerHisArgMetGluMetTrpLeuLeuLeuVal 23
|||||
37 AGGCACCTGG.....CATTACCAAGATGTGGCTGCTATTAGTAAC 74

23 lAlaTyrMetPheGln...ArgAsnValAsnSerValHisMetProThrL 39
|||||
75 AGTGTGCTATCTGCATTGGAGTGCACATGGCCCTATTGGAAAACCTGG 124

39 ysAlaValAspProGluAlaPheMetAsnIleSerGluIleIleGlnHis 55
|||||
125 GTCCCAAAAACCTCGAAGCAACATGAATGTTAGTCAGATGATACCTTAC 174

56 GlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspGlyTyrIl 72
|||||
175 TGGGGATATCCAAAGTCAGGAATATGAAGTTGTTACTGAAGATGGCTACAT 224

72 eLeuSerValAsnArgIleProArgGlyLeuValGlnProLysLysThrG 89
|||||
225 TCTGGGGTCTATAGAAATTCCTATGGGAAGAAATTCCTGAGAATATCG 274

89 lysArgProValValLeuLeuGlnHisGlyLeuValGlyGlyAlaSer 105
|||||
275 GCAAGAGACCTGTGCATATTTCAGCATGGTTTTCATTGTCATCAGCCACA 324

106 AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs 122
|||||
325 AACTGGATTACAAATCTGCCAAACACAGCCTGGCCTTCATTCTAGCAGA 374

122 pAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerA 139
|||||
375 TCGTGGCTATGATGTGGCTGGGACACAGTCGGAGGAATACATGGTCCC 424

139 rGlyHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
|||||
425 GGAATAATGTATACATTATTCACACAGACTCAGTTGATTTCTGGCCTTACG 474

```

```

156 TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLe 172
|||||
475 TTTGATGAATGGCTAAATATGACCTTCCAGCCACCACCTAGACTTCTATTGT 524

172 uGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSerGlnGlyT 189
|||||
525 ACAGAAACTGGACAGAGAAGATACACTATGTTGGTCACCTCTCAGGGCA 574

189 hrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
|||||
575 CCACATCGGTTTTATTGCCCTTTTCTACCAATCTGCTGCTGCTAAAAAA 624

206 IleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaL 222
|||||
625 ATCAAGAGGTTTTATGCAATAGCTCCAGTTGCTACTGTGAAGTATACAGA 674

222 sSerProGlyThrLysPheLeuLeuLeuProAspMetMetIleLysGlyL 239
|||||
675 AAGTCCCTTTAAAAAGATTTTCACTTATTCCTTAAGTTTCTTCAAGGTGA 724

239 euPheGlyLysLysGluPheLeu.....TyrGlnThrArgPheLeu 252
|||||
725 TATTGTGTAACAAAATGTTCTATGCCCCACAACTACTTAGATCAATTTCTT 774

253 ArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleCy 269
|||||
775 GGTACGGAAGTG.....TGCTCAGGGAGCTGTAGATCTTCTCTGTG 815

269 sSerAsnIleMetLeuLeuLeuGlyPheAsnThrAsnAsnMetAsnM 286
|||||
816 CAGCAACGCTTTATTCATCTTCTGTGGATTTCACAAAGAAACAACTTAATG 865

286 etSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVal 302
|||||
866 TGAGTCGCTTGTATCTAGGGCATTAATCCAGCAGGAACATCTACT 915

303 GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgAl 319
|||||
916 CAAGACCTTTTCCACTGGGCACAGCTTGTAAATCTGGGAAGCTTCAAGC 965

319 aPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnProTrp 336
|||||
966 CTATAACTGGGAAGTCCATTCACAGAACATGTTACACTACATCAGAAAA 1015

336 hrProValArgTyrArgValArgAspMetThrValProThrAlaMetTrp 352
|||||
1016 CGCCTCCCTACTATGATGTGTGACGCATGACCGTGCCCAATTGCAGTGTGG 1065

353 ThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuLe 369
|||||
1066 AACGGTGGCCATGACATCTGCTGCTGATCCCAAGATGTCGCAATGCTGCT 1115

369 uSerGluValThrAsnIleTyrHisLysAsnIleProGluTrpAlaH 386
|||||
1116 TCCAAACTCCCCAACCTTCTGTACCATAGGAGATTCTTCCCTCAATC 1165

386 lsValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGlu 402
|||||
1166 ACCTGGACTTCACTGGGCGATGGATGCGCTCAAGAGGTTTACAATGAG 1215

403 IleIleHisLeuMetGlnGlnGluThrAsn 413
|||||
1216 ATAGTTACCATGATGGCAGAAAGACTTAACAGAAAT 1248

seq_name: gb_htc:AK009479
seq_documentation_block:
LOCUS      AK009479      1343 bp      mRNA
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
            library, clone:231002K08, full insert sequence.
ACCESSION  AK009479
VERSION    AK009479.1 GI:12844303
KEYWORDS   CAP trapper.

```



```

/db_xref="taxon:10090"
/db_xref="MGD:MGI:1898391"
/db_xref="MGD:MGI:1914967"
/clone="2310022H17"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
52 1239
/seq="male"
/note="putative"
/codon_start=1
/protein_id="BAB26312.1"
/db_xref="GI:12844294"
/translation="MMLLVTSVLSAFGGAHLGFLGPKNPANNNVSOMITWGYG
SERVEVTEGGYILGVRIPIYKGNSENIGKRPVAYLOHGLIASATNATINLPNNLSA
FILADAGYDWLGNRSRNTWSRKNVYSPDSVEFWAFSPDEMAKYDLPATIDFIPVSKT
GQKIHVGSQGTIGTAFSPALAKRKIRFYALAPVATVVKYTESPFKISLIPK
FLKVIFGNKMFPHNYLDQFLGTECSRELLDLCSNALFIFCGDDKNNLNSREDFV
YLGHPAGTSTQDLFWAQLAKSGKIQAYMWGSPLOQMLHYNQKTPPYDVSAMTYPI
AVWNGGHDILADPDQVAMLLPKLPNLLYHKEILPYNHLDFINAMDAPQEVINEIVTM
AED"
BASE COUNT      395 a 288 c 283 g 379 t
ORIGIN

```

```

alignment_scores:
  Quality: 1138.50      Length: 411
  Ratio: 3.492          Gaps: 4
  Percent Similarity: 79.319 Percent Identity: 52.555

```

alignment_block:

US-09-333-159-47 x AK009473

Align seg 1/1 to: AK009473 from: 1 to: 1345

```

7  ArgGlnTrpIleValSerHisArgMetGluMetTrpLeuLeuLeuVal 23
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
34  AGGCACCTGG.....CACTACAAAGATGTGGCTGCTATTAGTAAC 71
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
23  lAlaTrpMetPheGln...ArgAsnValAsnSerValHisMetProThrL 39
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
72  AAGTGTGCTATCGCATTTGGAGGTGCACATGGCCCTATTGGAAAACTGG 121
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
39  yAlaValAspProGluAlaPheMetAsnIleSerGluIleIleGlnHis 55
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
122  GTCCCAAAACCTCGAAGCAACATGAAATGTTAGTCAGATGATACTTAC 171
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
56  GlnGlyTrpProCysGluGluTrpGluValAlaThrGluAspGlyTrpI 72
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
172  TGGGGATATCCAAAGTGAGGAATATGAAGTTGTTACTGAAGATGGCTAC 221
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
72  eLeuSerValAsnArgIleProArgGlyLeuValGlnProLysLysThrG 89
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
222  TCTGGGGTCTATAGAAATTCCTTATGGGAAGAAAAAATCTGAGATATCG 271
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
89  lYSerArgProValValLeuGlnHisGlyLeuValGlyGlyAlaSer 105
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
272  GCAAGAGACCTGTGGCATATTTCAGCATGGTGTGATTCATCAGCCACA 321
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
106  AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs 122
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
322  AACTGATTACAAATCTGCAACACACACACCTGGCCTTCATCTAGCAGA 371
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
122  pAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerA 139
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
372  TGTGCTGCTATGATGTGTGGTGGGAACAGTCGAGGGAATACATGTCCTC 421
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
139  rGlyHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
422  GGAATAATGTATATCTATCCACAGACTCAGTTGAATCTGGGCTTTCAGC 471
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
156  TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLe 172
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

472  TTTGATGAATGGCTAAATATATGACCTTCCAGCCACCATAGACTTCATTGT 521
172  uGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSerGlnGlyT 189
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
522  ACAGAAACTGGACAAGAGAATACACTATGTTGGTCACCTCTCAGGCA 571
189  hrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
572  CCACATATCGGTTTATTGCCCTTTTACCAATCCTGCTCTGGCTAAAAAA 621
206  lLysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLys 222
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
622  ATCAAGAGGTTTATGCAATAGTCCAGTGTGCTACTGCAAGATATACAGA 671
222  sSerProGlyThrLysPheLeuLeuProAspMetMetIleLysGlyL 239
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
672  AAGTCCCTTTTAAAGATTTTACATTATTCCTTAAGTTTCTCTCAAGGTGA 721
239  euPheGlyLysLysGluPheLeu.....TyrGlnThrArgPheLeu 252
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
722  TATTTGGTAAACAATGTTTCATGCCCCCACTACTTAGATCAATTCTT 771
253  ArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleCy 269
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
772  GGTACCGAAGTG.....TGCTCAGGGAGCTGCTAGATCTTCTCTG 812
269  sSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnAsnMetAsnM 286
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
813  CAGCAACGCTTATTCATCTTCTGTGGATTTCACAAGAAAAAATTAATG 862
286  etSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVal 302
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
863  TGAGTCGCTTTTCATGTGTATCTAGGGCATAATCCAGCAGGAACATCTACT 912
303  GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgAl 319
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
913  CAAGACCTTTTCCACTGGGCACAGCTTGCTAAATCTGGGAAGCTTCAAGC 962
319  apheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnProT 336
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
963  CTATACTGGGAAGTCCATTACAGAACATGTTACACTACAATCAGAAAA 1012
336  hrProValArgTyrArgValArgAspMetThrValProThrAlaMetTrp 352
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1013  CGCCTCCCTACTATGATGTGTCAGCATGACCGTGCCAAATTCAGGTGG 1062
353  ThrGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuLe 369
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1063  AACGGTGGCCATGACATCTGCTGATCCCAAGATGTCGCAATGCTGCT 1112
369  uSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAlaH 386
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1113  TCCCAAACTCCCAACCTTCTGTACCATAAGGAGATTTCCCTCAACATC 1162
386  lValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGlu 402
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1163  ACCTGGACTTCATCTGGCGATGGATGCCCTCAGAGGTTTACATGAG 1212
403  lIleIleHisLeuMetGlnGlnGluThrAsn 413
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1213  ATAGTTACCATGATGCGAGAAGAACTAACAGAAT 1245
seq_name: gb_hrc:AK009523
seq_documentation_block:
LOCUS      AK009523      1346 bp      mRNA
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
            library, clone:2310028I09, full insert sequence.
ACCESSION  AK009523
VERSION    AK009523.1  GI:12844369
KEYWORDS   CAP trapper.
SOURCE     Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library

```


	MUS MUSCULUS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
172	uGlnLysThrGlyGlnGluLeuValIleTyrTyrValGlyTyrSerGlnGlyt 189	Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
523	ACAGAAATCGCAAGAAGATACACTATCTGGTCACCTCAGGSCA 572	2 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
189	hrThrMetGlyPheIleAlaPheSerThrMetProgluLeuAlaGlnLys 205	2 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
573	CCACTATCGGTATTATGGCTTTTCTACCAATCCTGCTGCCTGAATAAAA 622	2 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
206	IleValMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLy 222	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
623	ATCAAGAGGTTTATGTCATTAAGTCCAGTGCTGCTGCTGCTGCTGCTG 672	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
222	sSerProGlyThrLysPheLeuLeuLeuLeuProAspMetHethLeyLyl 239	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
673	AAGTCCCTTTAAAGATTTCACTTATTCCTAAGTTTCTTCTCAAGGTGA 722	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
239	eUPheGlyLysLysGluPheLeu.....TyrGlnThrArgPheLeu 252	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
723	TATTTGGTACAANAATGTTTCATGCCACCACTACTTAGTCAATTTCTT 772	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
253	ArgGlnLeuValIleTyrLeuCysGlnValIleLeuLeuAsnGlnIleCy 269	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
773	GGTACGGAAGTG.....TGCTCACGGAGCTGCTAGATCTTCTCTG 813	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
269	sSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetAsnM 286	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
814	CACCAACGCTTTATTCATCTCTGTTGATTTGACCAAGAAACTTTAAATG 863	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
286	etSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVal 302	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
864	TGAGTCGCTTGATGTATCTAGGCATAATCCAGCAGGACATCTACT 913	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
303	GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgAl 319	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
914	CAAGACCTTTTCCACTGGGCACAGCTTGCTAAATCTGGGAAGCTTCAAGC 963	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
319	aPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnProt 336	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
964	CTAFACTGGGGAAGTCCATACAGGACATGTTACACTACAATCAGAAAA 1013	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
336	hrProValArgTyrArgValArgAspMetThrValProthrAlaMetTrp 352	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
1014	CGCTCCCTACTATGATGTGTCAGCATCACCGTCCCAATTCAGTGTGG 1063	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
353	ThrGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeule 369	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
1064	AACGGTGGCCATGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1113	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
369	uSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAlah 386	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
1114	TCCCAACTCCCCAACCTTCTGTACCATAAGAGAGATCTTCCTCAATC 1163	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
386	isValAspPheIleTrpGlyLeuAspAlaProHisArgMetYrAsnGlu 402	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
1164	ACCTGGACTTCATCTGGCGATGATGCGCTCAAGAGGTTTACATGAG 1213	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
403	IleIleHisLeuMetGlnGlnGluGluThrAsn 413	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
1214	ATAGTTACCATGATGCGCAAGAACTAACAGAAT 1246	3 (sites) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods Enzymol. 303, 19-44 (1999)
seq_name:	gb_hlc:AK009729	
seq_documentation_block:		
LOCUS	AK009729	1346 bp mRNA HTC 08-FEB-2001
DEFINITION	Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310040L03, full insert sequence.	
ACCESSION	AK009729	
VERSION	AK009729.1	GI:12844702
KEYWORDS	CAP trapper.	
SOURCE	Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library clone:2310040L03.	
FEATURES		
source		
Location/Qualifiers		
1..1346	/organism="Mus musculus"	
	/strain="C57BL/6J"	
	/db_xref="taxon:10090"	
	/db_xref="MGD:MGI:1898390"	
Host: SOLR.		
COMMENT		
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAAGGATCCCAAGCTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGAGAAGGATCCCAAGCTCTTTTTTTTTTTTNN 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.		
The RIKEN Genome Exploration Research Group Phase II Team and FANTOM Consortium.		
Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)		
5 (bases 1 to 1346)		
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Carninci,P., Fukuda,S., Fukushima,I., Furuno,M., Han		


```

/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
52..1239
CDS
/note="putative"
/codon_start=1
/protein_id="BAB26673.1"
/db_xref="GI:12845239"
/translation="MWLLVTVSVLSAFGGAHGLFGKLPKPNPEANNVNSQMTITWGYGYP
SSEYEVETDGYILGVIRIPYKNSENIGKRPVAYLQHLIASATNWITNLPNNSLA
FILADAGYVWLGNSRGNTWSKNNVYSPDSVEFWAFSDENAKYDLPATIDFIVQKT
COEKHYVCHSGMTIGTIGTAFSTNPALAKIKRFYALAPAVFKYTESPPKKISILPK
FLKLVIFGNKMFNPYLDOLFTEVCSRELLDLCSNALFIFCGFDRKLNLSRFV
YLGHPAGSTODLFEHQAOLSKGLQAYNWGSPLOKMLHYNOKTTPPYDYVSAMTVPI
AVMNGCHDLADPQDVAMLLPKLPNLLYHKEILPYNHLDFIWANDAPQEVYNEIVTM
AED"
BASE COUNT      396 a 290 c 284 g 378 t
ORIGIN

```

```

alignment_scores:
  Quality: 1138.50      Length: 411
  Ratio: 3.492          Gaps: 4
  Percent Similarity: 79.319  Percent Identity: 52.555

```

alignment_block:

US-09-333-159-47 x AK010058 ..

Align seg 1/1 to: AK010058 from: 1 to: 1348

```

7  ArgGlnTrpIleValSerHisArgMetGluMetTrpLeuLeuValLeuVa 23
   |||||
34 AGGCACCTGG.....CACTACAAGATGTGGCTATTAGTAAC 71
   |||||
23  lalaTyrMetPheGln...ArgAsnValAsnSerValHisMetProThrL 39
   |||||
72  AAGTGTGCTATCTGCATTGGAGTGGACATGGCGCTATTGGAAAACCTGG 121
   |||||
39  ysAlaValAspProGluAlaPheMetAsnIleSerGluIleGlnHis 55
   |||||
122 GTCCCAAAACCCCTGAAGCAACATGAATGTTAGTCAGATGATTAAC 171
   |||||
56  GlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspGlyTyrI 72
   |||||
172 TGGGGATATCCCAAGTGAAGATATGAAGTTGTTACTGAAGATGGCTACAT 221
   |||||
72  eLeuSerValAsnArgIleProArgGlyLeuValGlnProLysLysThrG 89
   |||||
222 TCTGGGGTCTATAGATTCCTTATGGGAGAAAANTCTGAGAAATATCG 271
   |||||
89  lySerArgProValValLeuLeuGlnHisGlyLeuValGlyGlyAlaSer 105
   |||||
272 GCAAGAGACCTGTGGCATAATTGTCAGCATGTTGATTGTCATCAGCCACA 321
   |||||
106 AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs 122
   |||||
322 AACTGGGATTACAAATCTGCCAAACACAGCCTGGCCCTTCATCTAGCAGA 371
   |||||
122  pAlaGlyPheAspValTyrMetGlyAsnSerArgGlyAsnAlaTrpSerA 139
   |||||
372 TGCTGGCTATGATGTGGTGGGAGACAGTCGAGGGNAATACATGGTCCC 421
   |||||
139  rGlySHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
   |||||
422 GGAATAATGTATCTATTCACCAGACTCAGTTGAATTCCTGGCTTCAGC 471
   |||||
156  TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLe 172
   |||||
472 TTTGATGAATGGGTAAATATGACCTCCAGCCACCACCATGACTTCATGT 521
   |||||
172  uGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSerGlnGlyT 189
   |||||

```

```

522 ACAGAAACTGGCACAAGAGAGATACACTATGTTGGTCACTCTCAGGGCA 571
189 hrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
   |||||
572 CCACATACGGTTTATTATGCTTTTCTACCAATCTGCTCTGGCTAAAAAA 621
206 IleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLy 222
   |||||
622 ATCAAGAGGTTTATGATTTAGTCCAGTCTGCTACTGTGAAGTATACAGA 671
222 sSerProGlyThrLysPheLeuLeuLeuProAspMetMetIleLysGlyL 239
   |||||
672 AAGTCCCTTTAAAAAGATTTTCACTATTCTTAAGTTTCTTCTCAAGGTGA 721
239 euPheGlyLysLysGluPheLeu.....TyrGlnThrArgPheLeu 252
   |||||
722 TATTTGTATACAAAATGTTTCATGCCCCCAACTACTAGATCAATTTCTT 771
253 ArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleCy 269
   |||||
772 GGTACGGAAGTG.....TGCTACGGGAGCTGCTAGATCTTCTCTG 812
269 sSerAsnIleMetLeuLeuLeuGlyPheAsnThrAsnAsnMetAsnM 286
   |||||
813 CAGCAACGCTTTATTCATCTCTGTGGATTTCACAAGAAAAAACTTAATG 862
286 etSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVal 302
   |||||
863 TGAGTCTCTTTGATCTGTATAGGCATATAATCCAGCAGGACATCTACT 912
303 GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgAl 319
   |||||
913 CAAGACCTTTTCCACTGGGCACAGCTTCTAAATCTGGGAAGCTTCAAGC 962
319 aPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnProT 336
   |||||
963 CTATAACTGGGGAAGTCCATTACAGACATGTTACACTACATCAGAAAA 1012
336 hrProValArgTyrArgValArgAspMetThrValProThrAlaMetTrp 352
   |||||
1013 CGCCTCCCTACTATGATGTGTGACGATGACGCGTGCCTCAATTCAGTGG 1062
353 ThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuLe 369
   |||||
1063 AACGTTGCCATGATCATCTCGCTGATCCCAAGATGTCGCAATCTGCT 1112
369 uSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAlaH 386
   |||||
1113 TCCCAAACTCCCAACCTTCTGTACCATAGGAGATTTCTCCCTACAATC 1162
386 lsValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGlu 402
   |||||
1163 ACCTGGACTTCATCTGGGCGATGGATGCGCTCAAGAGGTTTACAATGAG 1212
403 IleIleHisLeuMetGlnGlnGluThrAsn 413
   |||||
1213 ATAGTTACCATGGGCAGAGACTTAACAGAAAT 1245

```

seq_name: gb_hlc:AK009300

```

seq_documentation_block:
LOCUS      AK009300      1349 bp      mRNA      HTC      08-FEB-2001
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
            library, clone:2310011G18, full insert sequence.
ACCESSION  AK009300
VERSION    AK009300.1 GI:12844011
KEYWORDS   CAF trapper.
SOURCE     Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:2310011G18.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

```



```
189 hrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
573 CCACATCGGTTTATGCTTTTCTACCAATCTGCTGGCTAAAGAA 622
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
206 IleLysMetThrPheAlaLeuAlaProIleAlaThrValLysHisAlaLys 222
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
623 ATCAAGAGGTTTATGCAATAGCTCCAGTTGCTACTGTGAAGTATACAGA 672
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
222 sSerProGlyThrLysPheLeuLeuLeuProAspMetMetIleLysGlyL 239
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
673 AAGTCCCTTTAAAGATTTTCACTTATTCCTAAGTTTCTTCAAGGTGA 722
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
239 eupheGlyLysLysGluPheLeu.....TyrGlnThrArgPheLeu 252
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
723 TATTGTGTAACAAGTTTCATGCTCCACCACTACCTAGTACCAATTTCT 772
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
253 ArgGlnLeuValIleTyrLeuGlyGlnValIleLeuAspGlnIleCy 269
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
773 GGTACGGAAGTG.....TGCTCAGGGAGCTGCTAGATCTTCTCTG 813
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
269 sSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnMetAsnM 286
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
814 CAGCAACGCTTTATTCATCTCTGCTGATTTTGACAGAAACCTTAATG 863
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
286 etSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVal 302
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
864 TGAGTCGCTTGTGTGTATCTAGGCAATATCCAGCAGGAATCTTACT 913
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
303 GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgAl 319
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
914 CAGACCTTTTCCACTGGGACAGCTTGCTTAATCTGGGAGCTTCAAGC 963
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
319 aPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnProt 336
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
964 CTATACTGGGAGTCCATTACAGACATGTTTACATACATCAGAA 1013
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
336 hrProValArgTyrArgValArgAspMetThrValProThrAlaMetTrp 352
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1014 CGCTCCCTACTATGATGTCGACCCATGACCGTGCCCAATTTGCAAGTGG 1063
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
353 ThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuLe 369
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1064 AACGCTGGCATGACATCTGCGCTGATGCCCAAGATGTCGAATGCTGCT 1113
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
369 uSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAlaH 386
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1114 TCCCAAACTCCCAACCTTCTGTACCATAGGAGATCTTCCCTACAACT 1163
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
386 lsValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGlu 402
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1164 ACCTGGACTTCATCTGGCGCATGATGCGCTCAAGAGGTTTACAATGAG 1213
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
403 IleIleHisLeuMetGlnGlnGluThrAsn 413
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1214 ATATTACCATGATGGCAGAGACTACAGAAAT 1246
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
```

seq_name: gb_htc:AK009474

seq_documentation_block: 1349 bp mRNA HTC 08-FEB-2001
LOCUS AK009474 Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310022103, full insert sequence.

ACCESSION AK009474
VERSION AK009474.1 GI:12844295
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library clone:2310022103.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)

Carrinci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Methods Enzymol. 303, 19-44 (1999)
2 (sites)
Carrinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
3 (sites)
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, F., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
4 (sites)
The RIKEN Genome Exploration Research Group Phase II team and FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 (bases 1 to 1349)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Carrinci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, Y., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gscc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTNN 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGATCTCGAGTTTAATTAATTCCTCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

Host: SOLR.
Location/Qualifiers
1. .1349
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGI:MGI:1898389"
/db_xref="MGI:MGI:1914967"
/clone="2310022103"
/sex="male"
/tissue_type="tongue"

FEATURES
source

```

/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
53..1240
/notes="putative"
/codon_start=1
/protein_id="BAB26313.1"
/db_xref="GI:12844296"
/translaton="MWLLVTSVLSAFSGAHGLFGKLGPKNPANMNVSQMITVWGYP
SEEVYEDGVLGVYRIPYKKNSENIGKRPVAYLQHLGLIASATNWTNLPNNSLA
FILADAGYDVLGNSRGNTWRKNVYSPDSVEWAFSFDMAKYDILPATIDTFOKT
GOEKHYVHGSGGTIGFIAFSNPALAKIKRFPALAPVATKYTESPKKISLIPK
FLLVIVGNKRNPHNYLDOLFTEVCSRELLDLCSNALFICGFDKKNLVNSRFDV
YLGNPAGTSTQDLFWAQLAKSKLQAINWGSPLQNLHYNOKTPPYVDVSAMTVP
AVWNGGHDILADPDQVAMLLPKLPNLVYHREILFYNLHDFIAMDAPQEVYNEIVTM
AED"

```

```

BASE COUNT      397 a   290 c   284 g   378 t
ORIGIN

```

```

alignment_scores:
  Quality: 1138.50      Length: 411
  Ratio: 3.492          Gaps: 4
  Percent Similarity: 79.319  Percent Identity: 52.555

```

```

alignment_block:
US-09-333-159-47 x AK009474

```

```

Align seg 1/1 to: AK009474 from: 1 to: 1349

```

```

7  ArgGlnTrpIleValSerHisArgMetGluMetTrpLeuLeuIleLeuVa 23
35  AGGCACTGG.....CACTACAAGATGGTGGCTCTATTAGTAAC 72
23  lAlaTrpMetPheGln...ArgAsnValAsnSerValHisMetProThrL 39
73  AAGTGTGCTATCTGCATTGTGGAGGTGCACATGGCCTATTTGGAAACTGG 122
39  ySAlaValaspProGluAlaPheMetAsnIleSerGluIleleleGlnHis 55
123  GTCCCAAAACCTGTAAGCAACATGAATGTTAGTCAGATGATGAACCTTAC 172
56  GlnGlyTyrProCysGluGluTyrGluValAlaThrGluaspGlyTyril 72
173  TGGGGATATCCAAGTGAAGAAATGAAGTTGTACTGAAGATGGCTACAT 222
72  eLeuSerValAsnArgIleProArgGlyLeuValGlnProLysLysThrG 89
223  TCTGGGGTCTATAGAATTCCTTATGGGAAGAAATTCAGAAATATCG 272
89  lYSerArgProValValLeuLeuGlnHisGlyLeuValGlyAlaSer 105
273  GCAAGAGACCTGTGGCATATTTGGAGCATGTTGATTGCATCAGCCACA 322
106  AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs 122
323  ACTGGATTACAATCTGCCAACACACAGCCTGGCCTTCATTCTAGCAGA 372
122  pAlaGlyPheaspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerA 139
373  TGTGTGCTATGATGTGTGGCTGGGAACAGTCGAGGGAATACATGTCCTC 422
139  rGlyHisLysThrLeuSerIleaspGlnaspGluPheTrpAlaPheSer 155
423  GGAAAAATATATATCTATCTATTCACCACTCAGTTGAAATTCGGGCTTC 472
156  TyrAspGluMetAlaArgPheaspLeuProAlaValIleAsnPheIleLe 172
473  TTTGATGAATGGCTAAATATGACCTTCAGCCACCACATAGACTTCATGT 522
172  uGlnLysThrGlnGluLysIleTyrTrpValGlyTyrSerGlnGlyT 189
523  ACAGAAACTGGCAAGAGAAGATACACTATGTTGGTCACTCTCAGGCA 572

```

```

189  hrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
573  CCACATATCGTTTATTGCTTTTCTACCAATCCTCTCGCTGCTAAATAA 622
206  IleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAlaLy 222
623  ATCAAGAGGTTTTATGCATTAGCTCCAGTTGCTACTGTGAAGTATACAG 672
222  sSerProGlyThrLysPheLeuLeuLeuProaspMetMetIleLysGlyL 239
673  AAGTCCCTTTAAAGAGATTCTACTTATCTTAAGTTTCTTCAAGGTGA 722
239  euPheGlyLysGluPheLeu.....TyrGlnThrArgPheLeu 252
723  TATTGGTAACAAATGTTTCATGCCCCCACTACTTAGTCAATTTCTT 772
253  ArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleCy 269
773  GGTACGGGAAGTG.....TGCTACGGGAGCTGCTAGATCTTCTCTG 813
269  sSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnAsnMetAsnM 286
814  CAGCAACGCTTTATTCATCTCTGTGGATTGACAAAGAAAACTTAAATG 863
286  etSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVal 302
864  TGAGTCGCTTTCATGTGATCTAGGCATATATCCACGACGACATCTACT 913
303  GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgAl 319
914  CAAGACCTTTTCCACTGGGCACAGCTTGTAAATCTGGGAAGCTTCAAGC 963
319  aPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 336
964  CTATACCTGGGAAGTCCATCTACAGACATGTACACTACAATCAGAAAA 1013
336  hrProValArgTyrArgValArgAspMetThrValProThrAlaMetTrp 352
1014  CGCCTCCCTACTATGATGTGCAGCATGACCGTGCCAATTCGAGTGTGG 1063
353  ThrGlyGlyGlnAspTrpLeuSerAsnProGluaspValLysMetLeuLe 369
1064  AACGTGGCCATGACATCTGCTGTCATCCCAAGATGTCCGAATCTGCT 1113
369  uSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAlaH 386
1114  TCCCAACTCCCCACCTCTCTACATAAGGAGATCTTCCCTACAATC 1163
386  lSValAspPheIleTrpGlyLeuaspAlaProHisArgMetTyrAsnGlu 402
1164  ACCTGGACTTCATCTGGCGATGATGCGCTCAGAGGTTTACATATGAG 1213
403  IleIleHisLeuMetGlnGlnGluThrAsn 413
1214  ATAGTTACCATGATGGCAGAGACTAACAGAAT 1246
seq_name: gb_Htc:AK009546

```

```

seq_documentation_block:

```

```

LOCUS      AK009546      1349 bp      mRNA      HTC      08-FEB-2001
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
            library, clone:2310030L19, full insert sequence.
ACCESSION  AK009546
VERSION    AK009546.1  GI:12844406
KEYWORDS   CAP trapper.
SOURCE     Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
            clone_119:RIKEN full-length enriched mouse cDNA library
            Clone:2310030L19.

```

```

ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

```

```

REFERENCE  1 (sites)
AUTHORS   Carrinci,P. and Hayashizaki,Y.

```


/cqn2_6/ptodata/1/ina/5B_COMB.seq:US-08-600-908A-11 +	82.50	140.26	3.39
/cqn2_6/ptodata/1/ina/6A_COMB.seq:US-08-683-838A-11 +	82.50	140.26	3.39
/cqn2_6/ptodata/1/ina/5B_COMB.seq:US-08-483-101-1 +	82.50	122.45	33.26
/cqn2_6/ptodata/1/ina/5A_COMB.seq:US-08-348-891A-1 +	82.50	109.43	176.52
/cqn2_6/ptodata/1/ina/5A_COMB.seq:US-08-905-817-1 +	82.50	109.43	176.52

seq_name: /cqn2_6/ptodata/1/1na/5A_COMB.seq:US-08-227-108-2

seq_documentation_block:
; Sequence 2, Application US/08227108
Patent No. 5807,26

CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

score_list:	Sequence	Strd Orig	ZScore	EScore Len	Documentation	
	/cn2.6/ptodata/1/lna/5A.COMB	seq:US-08-227-108-2	1141.50	2461.52	1.7e-129	1137
	/cn2.6/ptodata/1/lna/5B.COMB	seq:US-09-073-674-2	1141.50	2461.52	1.7e-129	1137
	/cn2.6/ptodata/1/lna/5A.COMB	seq:US-08-227-108-4	1141.50	2461.49	1.7e-129	1140
	/cn2.6/ptodata/1/lna/5B.COMB	seq:US-09-073-674-4	1141.50	2461.49	1.7e-129	1140
	/cn2.6/ptodata/1/lna/5A.COMB	seq:US-08-227-108-6	1141.50	2461.42	1.7e-129	1146
	/cn2.6/ptodata/1/lna/5B.COMB	seq:US-09-073-674-6	1141.50	2461.42	1.7e-129	1146
	/cn2.6/ptodata/1/lna/5A.COMB	seq:US-08-227-108-1	1141.50	2457.71	2.8e-129	1528
	/cn2.6/ptodata/1/lna/5B.COMB	seq:US-09-073-674-1	1141.50	2457.71	2.8e-129	1528
	/cn2.6/ptodata/1/lna/6A.COMB	seq:US-09-336-952-12	271.50	510.09	8.5e-21	43676
	/cn2.6/ptodata/1/lna/5A.COMB	seq:US-08-751-782-2	155.00	326.10	1.5e-10	178
	/cn2.6/ptodata/1/lna/5B.COMB	seq:US-08-935-171-2	155.00	326.10	1.5e-10	178
	/cn2.6/ptodata/1/lna/5A.COMB	seq:US-08-484-105-21	105.00	187.71	0.0077	1676
	/cn2.6/ptodata/1/lna/5B.COMB	seq:US-08-484-106-21	105.00	187.71	0.0077	1676
	/cn2.6/ptodata/1/lna/5A.COMB	seq:US-08-602-359A-24	102.00	187.29	0.0081	1041
	/cn2.6/ptodata/1/lna/5A.COMB	seq:US-08-232-519-1	98.50	178.16	0.0262	1167
	/cn2.6/ptodata/1/lna/5B.COMB	seq:US-08-436-936-1	98.50	178.16	0.0262	1167
	/cn2.6/ptodata/1/lna/6B.COMB	seq:US-09-115-446-1	92.00	153.99	0.5820	2521
	/cn2.6/ptodata/1/lna/6B.COMB	seq:US-09-115-446-5	92.00	153.96	0.5839	2526
	/cn2.6/ptodata/1/lna/5A.COMB	seq:US-08-471-119A-1	92.00	116.26	73.51	46899
	/cn2.6/ptodata/1/lna/5A.COMB	seq:US-08-080-386-3	91.00	153.79	0.5967	2160
	/cn2.6/ptodata/1/lna/5B.COMB	seq:US-08-390-000A-3	91.00	153.79	0.5967	2160
	/cn2.6/ptodata/1/lna/6A.COMB	seq:US-08-742-195-102	91.00	116.02	75.82	40328
	/cn2.6/ptodata/1/lna/5B.COMB	seq:US-08-968-542C-1	89.00	136.17	5.72	6027
	/cn2.6/ptodata/1/lna/5B.COMB	seq:US-08-755-559-2	85.00	140.54	3.27	2180
	/cn2.6/ptodata/1/lna/6A.COMB	seq:US-09-210-474-2	85.00	140.54	3.27	2180
	/cn2.6/ptodata/1/lna/5B.COMB	seq:US-08-602-359A-30	84.50	149.29	1.06	1017
	/cn2.6/ptodata/1/lna/5B.COMB	seq:US-08-759-945-1	84.50	137.05	5.11	2625
	/cn2.6/ptodata/1/lna/5A.COMB	seq:US-07-707-367-1	84.50	133.78	7.78	3393
	/cn2.6/ptodata/1/lna/6B.COMB	seq:US-09-386-607-1	84.00	138.47	4.26	2160
	/cn2.6/ptodata/1/lna/5A.COMB	seq:US-07-882-339-4	82.50	141.46	2.90	1329
	/cn2.6/ptodata/1/lna/5A.COMB	seq:US-08-183-213-4	82.50	141.46	2.90	1329
	/cn2.6/ptodata/1/lna/5A.COMB	seq:US-08-446-803-4	82.50	140.29	3.37	1455
	/cn2.6/ptodata/1/lna/5B.COMB	seq:US-08-861-837-4	82.50	140.29	3.37	1455
	/cn2.6/ptodata/1/lna/6A.COMB	seq:US-08-602-656-4	82.50	140.29	3.37	1455
	/cn2.6/ptodata/1/lna/6B.COMB	seq:US-09-170-670-9	82.50	140.29	3.37	1455
	/cn2.6/ptodata/1/lna/6B.COMB	seq:US-09-170-670-14	82.50	140.29	3.37	1455
	/cn2.6/ptodata/1/lna/6B.COMB	seq:US-09-193-068-9	82.50	140.29	3.37	1455
	/cn2.6/ptodata/1/lna/6B.COMB	seq:US-09-193-068-13	82.50	140.29	3.37	1455
	/cn2.6/ptodata/1/lna/6B.COMB	seq:US-09-183-412-9	82.50	140.29	3.37	1455
	/cn2.6/ptodata/1/lna/6B.COMB	seq:US-09-183-412-13	82.50	140.29	3.37	1455

CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Fanucci, Allan A.
 REGISTRATION NUMBER: 30,256
 REFERENCE/DOCKET NUMBER: 7620-033
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 312 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1137 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 FEATURE:
 NAME/KEY: C

us-09-333-159-47.p2n.in1

110 GTTATACCTTGGGATCGACAGATTCCTTATGGGAGAAATTCAGAG 159
87 LysThrGlySerArgProValValLeuLeuGlnHisGlyLeuValGlyG1 103
160 AATATAGCCGAGACCTGTTCATTTTCACACAGGTTTGTCTCCATC 209
103 yalaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheLeu 120
210 AGCCACAACTGGATCTCCACCTGCCACACAGCCTGGCTTCATC 259
120 euAlaAspIleGlyPheAsnValTrpMetGlyAsnSerArgGlyAsnAla 136
260 TGCCGACCCGCGGTACGAGCTGGCTGGGACACAGCAGGCGCACACC 309
137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
310 TGGCCACAGAGAACTCTACTACTCGCCGACCTCGCTGCAATTCCTGGC 359
153 aPheSerTrpAspGluMetAlaArgPheAspLeuProAlaValIleAsn 170
360 TTTCAGCTTACAGAGATGGCTAAATAGACCTTCCGCCACCATGTGACT 409
170 heileLeuGlnLysThrGlyClnGlyLysIleTyrrValGlyTyrrSer 186
410 TCATCTTGAAGAAACGGGACAGGACAGCAAGCTACATACCTTGGCCATTC 459
187 GlnGlyThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
460 CAGGCAACCATCTGTTTCATCGCTTTCACCAATCCCAAGCTGCC 509
203 aGlnLysIleLysMetTyrrPheAlaLeuAlaProIleAlaThrValLysH 220
510 GAAACGATCAAAACCTTCTATGATAGCTTCCGCTGCCACCTGAGT 559
220 IsalaLysSerProGlyThrLysPheLeuLeuProAspMetMetIle 236
560 ACACGAAACCTGTAAACAACTCATCTCGCTCCCTCGTCTCTTC 609
237 LysGlyLeuPheGlyLysGluPheLeuTyrrGlnThrArgPheLeuAr 253
610 AAGCTTATATTGAAACAAATATTC...TACCACACCACTCTTTTGA 656
253 gGln...LeuValIleTyrrLeuCysGlyGlnValIleLeuAspGlnIleC 269
657 TCAATTTCTGCCACGAGGTATCTCCCGGAGAGCGGTGATCTCTCT 706
269 ySerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnAsnMetAsn 285
707 GCAGCAACCCCTTTATCATCTTTCGATTTTCACACTGACACTGA 756
286 MetSerArgAlaSerValTyrrAlaAlaHisThrLeuAlaGlyThrSerVa 302
757 ATGAGTCTGGTGTGTATCTGTACATAATCCAGGAGCAATCGT 806
302 IGlAsnIleLeuHisTrpSerClnAlaValAsnSerGlyGlyLeuArgA 319
807 TCAGACGCTGCTCATCTGCTCCAGGCTGTAGTCTGGGAAGTCCAG 856
319 laPheAspTrpGlySerGlnThrLysAsnLeuGluLysCysAsnGlnPro 335
857 CTTTGTAGTGGGAAGCCAGCTTCAGAACATGGCTCATCATCATCAG 906
336 ThrProValArgTyrrAlaArgAspMetThrValProThrAlaMetTr 352
907 ATCCCTCTCTATCAACCTGACAGACATCATGTGCCAATCCGACGTG 372
352 pThrGlyClnGlnAspTrpLeuSerAsnProGluAspValLysMetLeu 385
957 GAACGGTGGCAGCACTGTGCGGACCTTCAGATGTGACCTTTTC 1006
957 eUerGluValThrAsnLeuIleTyrrHisLysAsnIleProGluTrpAla 1056
1006 eUerGluValThrAsnLeuIleTyrrHisLysAsnIleProGluTrpAla 1056

1007 TTTCAGCTTCCCAATCTCATTTACCACAGAGATTCCTCTTACAAT 1056
386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrrAsnG1 402
1057 CACTTGAGCTTATCTGGGCCATGATGCCCTCAAGCGGTTTACAATCA 1106
402 uileIleHisLeuMet 407
1107 AATGTTTTCATGATG 1122
seq_name: /cgn2_6/p/odata/1/ina/5B_COMB_seq:US-09-073-674-2
seq_documentation_block:
Sequence 2, Application US/09073674
Patent No. 5998189
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claire
TITLE OF INVENTION: Jean-Louis
NUMBER OF SEQUENCES: Recombinant Dog Gastric Lipase
CORRESPONDENCE ADDRESS: 21
ADDRESS: Warner-Lambert Company
STREET: 2800 Plymouth Road
CITY: Ann Arbor
STATE: Michigan
COUNTRY: U.S.A.
ZIP: 48105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/09/073,674
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crissey, Todd M.
REGISTRATION NUMBER: 37,807
TELEPHONE: 734 622-7530
TELEFAX: 734 622-1553
INFORMATION FOR SEQ ID NO. 2:
LENGTH: 1137 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
NAME/KEY: CDS
LOCATION: 1..1137
US-09-073-674-2
alignment_scores:
Quality: 1141.50
Ratio: 3.624
Percent Similarity: 84.677
Length: 372
alignment_block:
US-09-333-159-47 x US-09-073-674-2
Align seg 1/1 to: US-09-073-674-2
37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIle1 53
19 CCCACA.....AACCTCAAGG.....
957 GAACGGTGGCAGCACTTGTGCGGACCTTCAGATGTGACCTTTTC 1006
369 eUerGluValThrAsnLeuIleTyrrHisLysAsnIleProGluTrpAla 385
1006 eUerGluValThrAsnLeuIleTyrrHisLysAsnIleProGluTrpAla 1056

e Sep 4 10:30:00

Percent Similarity: 83.171 Percent Identity: 60.732

alignment_block:
US-09-333-159-47 x HSLAL ..

Align seg 1/1 to: HSLAL from: 1 to: 2626

```

2  LeuGluThrLeuSerArgGlnTrpIleValSerHisArgMetGluMetTr 18
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
113 CTCGAGACGCGCGCGGCAG.....GACAGCTCCAGAAATGAAATATCGG 156

18 pLeuLeuIleLeuValalaIleValMetPheGlnArgAsnValAsnSerValH 35
:|||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
157 GTTCTTGGGGTGGTGGTCTGTTGGTGTCTCTCGACCCCTGCATTCTCTGAGG 206

35 IsMetProThrLys.....AlaValAspProGluAlaPheMetAsnIle 49
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
207 GGTCTGGAGGAACTACAGACTGTGGATCTCGAAACAACATGAATATGG 256

50 SerGluIleLeuGlnHisGlnGlyTyrProCysGluGluTyrGluValAl 66
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
257 AGTCAAAATATCTCTTACTGGGATTCCTCTAGTAGGAAATACCTACTTGA 306

66 aThrGluAspGlyTyrIleLeuSerValAsnArgIleProArgGlyLeuV 83
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
307 GACAGAAGATGGATATATCTGTGCCTTACCCGAATTCCTCATGGGAGGA 356

83 alGlnProLysLysThrGlySerArgProValValLeuLeuGlnHisGly 99
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
357 AGACCATCTCGAAGGTCCCAACCACTGTCTCTCGCAACATGGC 406

100 LeuValGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLe 116
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
407 TTGCTGGCAGATTCTAGTAACGTGGTCACAAACCTTCCCAACAGCAGCCT 456

116 uClyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGlyAsnSerA 133
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
457 GGGCTTCATTCTTGTCTGATCTGGTTTGCAGTGTGGATGGCAACAGCA 506

133 rgGlyAsnAlaTrpSerArgLysHisLysThrLeuSerIleAspGlnAsp 149
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
507 GAGGAATACCTGGTCTCGGAACACATAAGACACTCTCAGTTCTTCAGGAT 556

150 GluPheTrpAlaPheSerTyrAspGluMetAlaArgPheAspLeuProAl 166
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
557 GAATCTGGGCTTTCAGTTATGATGAGATGGCAAAATATGACCTACCAGC 606

166 aValIleAsnPheIleLeuGlnLysThrGlyGlnGluLysIleTyrTyrV 183
| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
607 TTCATTAACTTCATTCTGAATAAACTGGCCCAAGCAAGTGTATTATG 656

183 alGlyTyrSerGlnGlyThrThrMetGlyPheIleAlaPheSerThrMet 199
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
657 TGGGTCTATTCTCAAGGCACCACTATAGGTTTTATACATTTTCCACAGATC 706

200 ProGluLeuAlaGlnLysIleLysMetTyrPheAlaLeuAlaProIleAl 216
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
707 CCTGAGCTGGCTAAAGGATTAATGTTTTTGGCCTGGGCTCTGTGGC 756

216 aThrValLysHisAlaLysSerProGlyThrLysPheLeuLeuLeuProA 233
|:|||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
757 TTCGCTCGCCTCTGTACTACGCCCTATGGCCAAATTAGGACGATTACCAG 806

233 spMetMetIleLysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThr 249
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
807 ATCATCTCAATAGGACTTATTTGGAGACAAAGAATTTCTTCCCCAGAGT 856

250 ArgPheLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAs 266
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
857 GCGTTTTTGAAGTGGCTACCCAGGTTTGCACTCATGTCATCTACTGAA 906

266 pGlnIleCysSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnA 283
:||||:||||:||||: ||||| ||||| ||||| ||||| ||||| |||||
```

